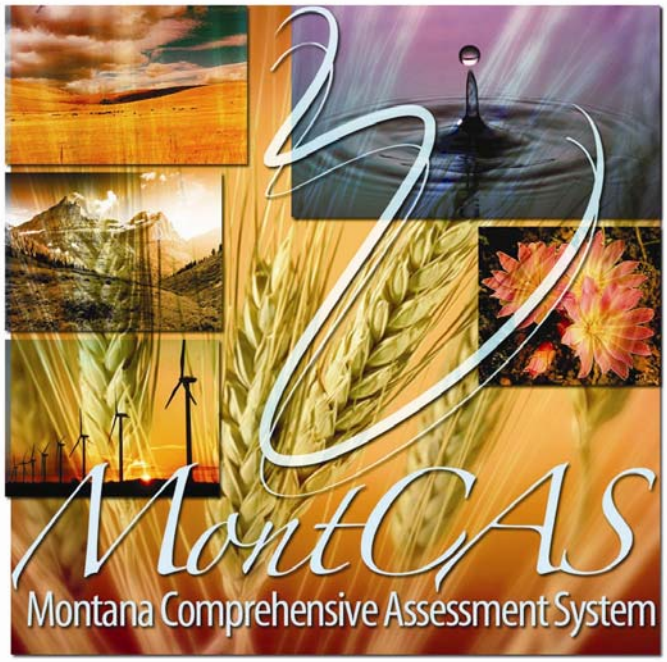


MontCAS
Criterion-Referenced Test
MONTANA CRT
2009–10
Technical Report

Measured Progress

100 EDUCATION WAY, DOVER, NH 03820 www.measuredprogress.org

TABLE OF CONTENTS

CHAPTER 1. OVERVIEW OF THE MONTANA CRITERION-REFERENCED TEST

1.1 PURPOSE OF THE ASSESSMENT SYSTEM

The Montana Criterion-Referenced Test (CRT) was developed in accordance with the following federal laws: Title 1 of the Elementary and Secondary Education Act (ESEA) of 1994, P. L. 103–382 and the No Child Left Behind Act (NCLB) of 2001.

In the spring of 2010, Montana students in grades 3 through 8 and 10 participated in the MontCAS Criterion-Referenced Test (Montana CRT) in reading, mathematics, and science. The purpose of this assessment is to measure students' achievement as articulated by Montana content standards and grade-level expectations.

All Montana students enrolled in accredited schools are required to participate in either the Montana CRT or the Montana CRT-Alternate. The vast majority of students participate in the CRT using standard administration procedures. However, an array of standard accommodations is available to any student, with or without disabilities, when such accommodations are necessary to allow the student to demonstrate his/her skills and competencies. Standard accommodations are not considered to change the constructs being measured and may be provided to students as necessary for any or all of the reading, mathematics, or science portions of the assessment. Students' tests are scored the same way whether they take the test using standard accommodations or not.

In addition to standard accommodations, other accommodations for the Montana CRT are available to students when specified in their IEPs, 504 plans, or LEP plans. These other accommodations are referred to as nonstandard accommodations; because they are considered to alter the constructs being measured, they do affect a student's score on the CRT. When a nonstandard accommodation is used, the student's score is reported as the lowest possible for that content area (e.g., a scaled score of 200 will fall into the Novice performance level). Nonstandard accommodations may be provided in reading, mathematics, or science, as dictated by the student's IEP, 504 plan, or LEP plan.

A small percentage of students participate in the statewide assessment program by taking the CRT-Alternate. Students with significant cognitive disabilities, who are working toward alternate academic achievement standards, as documented in their Individualized Education Program (IEP), are eligible to take the CRT-Alternate. Technical characteristics of the CRT-Alternate program are described in a companion technical report.

Montana grade-content CRT instruments are based on and aligned to Montana's content standards, benchmarks, and grade-level expectations in reading, mathematics, and science. Montana educators worked with the Montana Office of Public Instruction (OPI) and Measured Progress to develop test items that assess how well students have met Montana grade-level expectations for each content area. In addition, Northwest

Regional Educational Laboratory (NWREL performed an independent alignment study for mathematics and reading in 2006 and for science in 2007. NWREL's alignment studies can be found on the OPI's Web site at www.opi.mt.gov/assessment.

Montana CRT scores are intended to be useful indicators of the extent to which students have mastered material outlined in Montana reading, mathematics, and science content standards, benchmarks, and grade-level expectations. Each student's Montana CRT score should be used as part of a body of evidence regarding mastery and should not be used in isolation to make high-stakes decisions. Montana CRT scores are more reliable indicators of program success when aggregated to school, system, or state levels, particularly when monitored over the course of several years.

1.2 PURPOSE OF THIS REPORT

This report describes technical aspects of the Montana CRT in an effort to contribute to the accumulation of validity evidence to support Montana CRT score interpretations. Because the interpretations of test scores, not the test itself, are evaluated for validity, this report presents documentation to substantiate intended interpretations (American Educational Research Association [AERA], American Psychological Association & National Council on Measurement in Education, 1999). Subsequent chapters of this report discuss test development, test alignment, test administration, scoring, equating, item analyses, reliability, scaled scores, performance levels, and reporting. Each of these topics contributes important information toward establishing the validity of the assessment program. Note, however, that this report does not include certain aspects of a comprehensive validity argument that could also be important to consider when drawing conclusions about validity (e.g., additional sources of validity evidence might speak to the extent to which Montana CRT scores converge with other measures of the same or similar constructs and diverge from measures of different constructs, consequences that arise from scores at the student, school, district, and state levels).

Historically, some parts of technical reports may have been used by educated laypersons, but the intended audience was experts in psychometrics and educational research. This edition of the Montana CRT Technical Report attempts to make information more accessible to educated laypersons by providing more thorough descriptions of general categories of information. In making some information more accessible, we have purposely preserved the depth of technical information provided. The reader will find that some discussions and tables continue to require a working knowledge of measurement concepts, such as "reliability" and "validity," and statistical concepts, such as "correlation" and "central tendency." To fully understand some of the data presented, the reader will have to possess a basic understanding of advanced topics in measurement and statistics.

CHAPTER 2. CURRENT YEAR UPDATES

The state decided to collect certain testing demographic information through the state student information system (AIM). Thus, the bubbles for former LEP, Alt, NSAY, and NDAY on the answer booklet were removed and these data were supplied to Measured Progress from the OPI.

In 2010 the number of field test forms was reduced from eight forms to four forms.

In 2010 a practice question has been added to each content section.

CHAPTER 3. ASSESSMENT AND TEST DEVELOPMENT PROCESS

3.1 TEST SPECIFICATIONS

3.1.1 Criterion-Referenced Test

Items on the Montana Criterion-Referenced Test (CRT) are developed specifically for Montana and are directly linked to Montana's content standards. These content standards are the basis for the reporting categories developed for each content area and are used to help guide the development of test items. No other content or process is subject to statewide assessment. An item may address part, all, or several of the benchmarks within a standard.

3.1.2 Item Types

Montana educators and students are familiar with the types of items used in the assessment program. The types of items and their functions are described below:

- **Multiple-choice (MC)** items are used to provide breadth of coverage within a content area. Because they require no more than a minute for most students to answer, multiple-choice items make efficient use of limited testing time and allow for coverage of a wide range of knowledge and skills.
- **Short-answer (SA**) mathematics items are used to assess students' skills and abilities to work with brief, well-structured problems that have one or a very limited number of solutions (e.g., mathematical computations). Short-answer items require approximately two minutes for most students to answer. The advantage of this type of item is that it requires students to demonstrate knowledge and skills by generating, rather than merely selecting, an answer.
- **Constructed-response (CR)** items typically require students to use higher-order thinking skills—evaluation, analysis, summarization, and so on—to construct satisfactory responses. Constructed-response items take most students approximately five to ten minutes to complete. Note that the use of released Montana CRT items to prepare students to respond to constructed-response items is appropriate and encouraged.

3.1.3 Description of Test Design

The Montana CRT is structured using both *common* and *field-test* items. Common items are taken by all students in a given grade level. Student scores are based only on common items. In addition, field-test items are divided among the four forms of the test for each grade level. Each student takes only one form of the test and therefore answers a fraction of the field-test items. Field-test items are not identifiable to test takers and have a negligible impact on testing time. Because all students participate in the field test, it

provides the sample size (750–1500 students per item) needed to produce reliable data that can be used to inform item selection for future tests.

3.2 READING TEST SPECIFICATIONS

3.2.1 Standards

The test specifications/blueprint for reading is based on Montana's reading content standards, which identify five Montana content standards that apply specifically to reading and reading comprehension. Those content standards are listed below:

- **Reading Standard 1:** Students construct meaning as they comprehend, interpret, and respond to what they read.
- **Reading Standard 2:** Students apply a range of skills and strategies to reading.
- **Reading Standard 3:** Students set goals, and monitor and evaluate their reading progress. (This standard cannot be measured with a traditional paper-pencil test.)
- **Reading Standard 4:** Students select, read, and respond to print and non-print materials for a variety of purposes.
- **Reading Standard 5:** Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audience.

3.2.2 Item Types

The Montana CRT in reading includes a mix of multiple-choice and constructed-response items. Constructed-response items require students to write answers consisting of one or more paragraphs. Each type of item is worth a specific number of points in the student's total reading score, as shown in Table 3-1.

Table 3-1. 2009–10 MontCAS: Item Types

Item type	*Possible score points*
MC	0 or 1
CR	1, 2, 3, or 4

MC = multiple-choice; CR = constructed-response

3.2.3 Test Design

Table 3-2 shows the numbers of multiple-choice and constructed-response items for grades 3 through 8 and 10.

Table 3-2. 2009–10 MontCAS: Common Reading Items—Grades 3–8 and 10

				Total	
Grade	Session 1	Session 2	Session 3	MC	CRs
3–8	19 MC, 1 CR	14 MC	19 MC, 1 CR	52	2
10	19 MC, 1 CR	14 MC	19 MC, 1 CR	52	2

MC = multiple-choice; CR = constructed-response

3.2.4 Blueprints (Distribution of Points Across Standards)

Table 3-3 shows the distribution of points across content standards.

Table 3-3. 2009–10 MontCAS: Reading Specifications/Blueprint Grades 3–8 and 10

Number of points for the common (scored) test:	*52 MC items + 2 CR items = 60 points*						
Percent point distribution by content standard*							
Content Standards	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Standard 1	34%	34%	34%	34%	34%	34%	25%
Standard 2	30%	30%	30%	30%	30%	30%	32%
Standard 3							
Standard 4	18%	18%	18%	18%	18%	18%	22%
Standard 5	18%	18%	18%	18%	18%	18%	22%
*Because percents are rounded to the nearest whole number, not all sums add to 100%. Note: Standard 3 cannot be measured with a traditional paper-pencil test.							
Target point distribution by content standard (acceptable range)							
Content Standards	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Standard 1	20 (18–22)	20 (18–22)	20 (18–22)	20 (18–22)	20 (18–22)	20 (18–22)	15 (13–17)
Standard 2	18 (16–20)	18 (16–20)	18 (16–20)	18 (16–20)	18 (16–20)	18 (16–20)	19 (17–21)
Standard 3							
Standard 4	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	13 (11–15)
Standard 5	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	11 (9–13)	13 (11–15)

MC = multiple-choice; CR = constructed-response
Four-point items: Each test contains two four-point constructed-response items. In any given year, the two items will measure two different standards. From year to year, those standards may change.
One-point items: The number of one-point items per content standard will vary from year to year depending on which two standards are measured by the four-point items. (The number of total points per standard falls within the acceptable range from year to year.)

3.2.5 Depth of Knowledge

Each item on the Montana CRT in reading is assigned a depth-of-knowledge (DOK) level. The depth-of-knowledge level reflects the complexity of mental processing students use to answer an item. Depth of knowledge is not synonymous with difficulty. Each of the levels is described below.

- **Level 1 (Recall).** This level requires students to receive or recite facts or to use simple skills or abilities. Items require only a shallow understanding of the text presented and often consist of

verbatim recall from text, slight paraphrasing of specific details from the text, or simple understanding of a single word or phrase.

- **Level 2 (Skill/Concept).** This level includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Inter-sentence analysis of inference is required. Some important concepts are covered, but not in a complex way. Standards and items at this level may include words such as summarize, interpret, infer, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed.
- **Level 3 (Strategic thinking).** Deep knowledge becomes a greater focus at Level 3. Students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Standards and items at Level 3 involve reasoning and planning. Students must be able to support their thinking. Items may involve abstract theme identification, inference across an entire passage, or students' application of prior knowledge. Items may also involve more superficial connections between texts.

3.2.6 Passage Types

Reading passages include both long and short texts selected from sources that students in each grade level would likely encounter in their classroom or in their independent reading. No passages were written specifically for the assessment but were instead collected from published works. Each passage is classified as described below.

- **Literary passages** are represented by a variety of genres—modern narratives; diary entries; drama; poetry; biographies; essays; excerpts from novels; short stories; and traditional narratives, such as fables, myths, and folktales.
- **Informational passages** are nonfiction and generally include two subgenres.
- **Content passages** are primarily informational and often deal with the areas of science and social studies. They are drawn from sources such as newspapers, magazines, and books.
- **Practical passages** are functional materials that instruct or advise the reader—for example, directions, reference tools, or reports.

The main difference among the passages used for grades 3 through 8 and 10 is their degree of complexity, which results from increasing levels of sophistication in language and concepts, as well as passage length. Measured Progress uses a variety of readability formulas to aid in the selection of passages appropriate for the intended audience. In addition, the grade-level expertise of Montana teachers contributes to the selection of passages for each grade level.

Items related to these passages require students to demonstrate their skills in both literal comprehension, where the answer is stated explicitly in the text, and inferential comprehension, where the answer is implied by the text and/or the text must be connected to relevant prior knowledge to determine an answer. Items focus on the reading skills reflected in the content standards and require students to use reading skills and strategies to answer correctly—for example, how to identify the author's principal purpose, such as to persuade, entertain, or inform—and to demonstrate their understanding of how words and images communicate to readers. Tables 3-4 and 3-5 depict passage distribution and length in grades 3 through 8 and 10.

Table 3-4. 2009–10 MontCAS:
Reading Passage Distribution Grades 3–8 and 10

Passage type	*Passage content*	*Percent of test*	*Point distribution*
Literary	Stories, poetry, and other forms of literature	50%	30 points
Informational	Content and practical passages	50%	30 points
	Total		60 points

Passage length	*Passage type*	*Percent of test*	*Point distribution*
Long	One literary or one informational per session	50%	30 points
Short	At least one literary and informational per session	50%	30 points
	Total		60 points

Table 3-5. 2009–10 MontCAS:
Approximate Length of Reading Passages

Grade	*Long passage (number of words)*	*Short passage (maximum word length)*
3	350–800	350
4	400–850	400
5	450–850	450
6	450–900	450
7	450–950	450
8	500–1,000	500
10	550–1,200	550

While every attempt is made to adhere to recommended grade-level word counts for long and short passages, the final decision to select a passage is based on extensive reviews by content experts and bias panels, careful analysis of the sophistication of language and complexity of concepts in the passage, and the readability of the passage.

3.3 MATHEMATICS TEST SPECIFICATIONS

3.3.1 Standards

The mathematics specifications/blueprint is based on Montana's mathematics content standards:

- Mathematics Standard 1: Problem Solving

- Mathematics Standard 2: Numbers and Operations
- Mathematics Standard 3: Algebra
- Mathematics Standard 4: Geometry
- Mathematics Standard 5: Measurement
- Mathematics Standard 6: Data Analysis, Probability, and Statistics
- Mathematics Standard 7: Patterns, Relations, and Functions

3.3.2 Item Types

The Montana CRT in mathematics includes multiple-choice, short-answer, and constructed-response items. Short-answer items require students to perform a computation or solve a simple problem. Constructed-response items are more complex, requiring 8–10 minutes of response time. Each type of item is worth a specific number of points in the student's total mathematics score, as shown in Table 3-6.

Table 3-6. 2009–10 MontCAS: Item Types

Item type	*Possible score points*
MC	0 or 1
SA	0 or 1
CR	0, 1, 2, 3, or 4

MC = multiple-choice; SA = short-answer; CR = constructed-response

3.3.3 Test Design

Table 3-7 summarizes the numbers and types of items that were used to construct the common portion of the Montana CRT in mathematics for 2009–10.

Table 3-7. 2009–10 MontCAS: Common Mathematics Items

Session	*Calculator*	*Number of items grades 3, 4, 5*	*Calculator*	*Number of items grades 6, 7, 8, 9, 10*
1	Not Allowed	18 MC 2 SA 1 CR	Not Allowed	14 MC 3 SA 1 CR
2	Not Allowed	19 MC 1 SA	Allowed	21 MC
3	Allowed	18 MC 1 CR	Allowed	20 MC 1 CR

MC = multiple-choice; SA = short-answer; CR = constructed-response

3.3.4 Blueprints (Distribution of Points Across Standards)

Table 3-8 shows the distribution of points across the content standards.

Table 3-8. 2009–10 MontCAS: Mathematics Specifications/Blueprint

*Raw score/percent point distribution by content strand)**							
Content Standards	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Problem Solving and Number and Operations	34%(22)	34%(22)	32%(21)	32%(21)	27%(18)	27%(18)	20%(13)
Algebra	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	16%(11)
Geometry	15%(10)	15%(10)	16%(11)	16%(11)	18%(12)	18%(12)	20%(13)
Measurement	15%(10)	15%(10)	12%(8)	12%(8)	12%(8)	12% (8)	12%(8)
Data Analysis, Probability, and Statistics	12%(8)	12%(8)	15%(10)	15%(10)	18%(12)	18%(12)	20%(13)
Patterns, Relations, and Functions	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)	12%(8)

*Because percents are rounded to the nearest whole number, not all sums add to 100%.

The mathematics test design consists of 55 multiple-choice items, three one-point short-answer items, and two four-point constructed-response items for 66 total points. There are two types of one-point items: multiple-choice and short-answer. The number of one-point items per strand will vary from year to year depending on which two strands are measured by the four-point items.

3.3.5 Depth of Knowledge

Each item on the Montana CRT in mathematics is assigned a depth-of-knowledge (DOK) level according to the cognitive demand of the item. Depth of knowledge is not synonymous with difficulty. The depth-of-knowledge level rates the complexity of the mental processing a student must use to solve a problem. Each of the levels is described below:

- **Level 1 (Recall).** This level requires the recall of a fact, definition, term, or simple procedure; the application of a formula; or the performance of a straight algorithmic procedure. Items at this level may require students to demonstrate a rote response.
- **Level 2 (Skill/Concept).** This level requires mental processing beyond that of a habitual response. These items often require students to make some decisions about how to approach a problem.
- **Level 3 (Strategic Thinking).** This level requires students to develop a plan or sequence of steps. These items are more complex and abstract than the items at the previous two levels. These items may also have more than one possible answer and may require students to use evidence, make conjectures, or justify their answers.

It is important that the Montana CRT in mathematics measure a range of depths of knowledge. Table 3-9 shows the percent and point ranges of the three depth-of-knowledge levels used on the CRT in mathematics.

Table 3-9. 2009–10 MontCAS: Depth-of-Knowledge (DOK) Percent and Distribution by Level

DOK level	*Percent range*	*Point range*
1	20% to 30%	13 to 20 points
2	60% to 75%	39 to 50 points
3	5% to 10%	4 to 8 points

3.3.6 Use of Calculators and References Sheets

Montana educators who helped develop the Montana CRT acknowledged the importance of mastering arithmetic algorithms. At the same time, they understood that the use of calculators is a necessary and important skill in society today. Calculators can save time and prevent error in the measurement of some higher-order thinking skills and allow students to do more sophisticated and intricate problems. For these reasons, calculators were permitted on some parts of the Montana CRT in mathematics and prohibited on other parts. (Students were allowed to use any calculator with which they were familiar.)

3.4 SCIENCE TEST SPECIFICATIONS

3.4.1 Standards

The science specifications are based on Montana's science content standards:

- **Science Standard 1:** Scientific Investigations—Students, through the inquiry process, demonstrate the ability to design, conduct, evaluate, and communicate results and reasonable conclusions of scientific investigations.
- **Science Standard 2:** Physical Science—Students, through the inquiry process, demonstrate knowledge of properties, forms, changes, and interactions of physical and chemical systems.
- **Science Standard 3:** Life Science—Students, through the inquiry process, demonstrate knowledge of characteristics, structures and function of living things, the process and diversity of life, and how living organisms interact with each other and their environment.
- **Science Standard 4:** Earth/Space Science—Students, through the inquiry process, demonstrate knowledge of the composition, structures, processes, and interactions of Earth's systems and other objects in space.
- **Science Standard 5:** Impact on Society—Students, through the inquiry process, understand how scientific knowledge and technological developments impact communities, cultures, and societies.
- **Science Standard 6:** Historical Development—Students understand historical developments in science and technology.

3.4.2 Item Types

The CRT in science includes multiple-choice and constructed-response items. Multiple-choice items require students to select the correct response from four choices, each item taking an average of one minute to answer. Constructed-response items are more involved, requiring 8–10 minutes of response time. Each type of item is worth a specific number of points in the student's total science score, as shown in Table 3-10.

Table 3-10. 2009–10 MontCAS: Item Types

Item type	*Possible score points*
MC	0 or 1
CR	0, 1, 2, 3, or 4

MC = multiple-choice; CR = constructed-response

3.4.3 Test Design

Table 3-11 summarizes the numbers and types of items that were used to compute student scores on the 2009–10 Montana CRT in science. Additionally, each test form had 13 multiple-choice matrix field-test items and one constructed-response field-test item that did not affect student scores.

Table 3-11. 2009–10 MontCAS: Science Items

Grades	*Session 1*	*Session 2*	*Session 3*	*TOTAL*	
				MC	CR
4, 8, and 10	18 MC, 1 CR	17 MC	18 MC, 1 CR	53	2

MC = multiple-choice; SA = short-answer; CR = constructed-response

3.4.4 Blueprints (Distribution of Points Across Standards)

Table 3-12 shows the distribution of points and item types across the content standards.

Table 3-12. 2009–10 MontCAS: Science Specifications/Blueprint

Percent point distribution by content standard	
Montana Standards	Grades 4, 8, and 10
1. Scientific Investigations	23%
2. Physical Science	23%
3. Life Science	23%
4. Earth/Space Science	23%
5. Impact on Society 6. Historical Development	8%

Point distribution by content standard	
Montana Standards	Grades 4, 8, and 10
1. Scientific Investigations	14
2. Physical Science	14
3. Life Science	14
4. Earth/Space Science	14

continued

Point distribution by content standard	
5. Impact on Society 6. Historical Development	5

The science test design consists of 53 multiple-choice items and two four-point constructed-response items for 61 total points. In any given year, the two constructed-response items will measure two different standards. From year to year, those standards may change.

3.4.5 Depth of Knowledge

Each item on the Montana CRT in science is assigned a depth-of-knowledge (DOK) level. The depth-of-knowledge level reflects the complexity of mental processing students use to answer an item. Depth of knowledge is not synonymous with difficulty. Each of the levels is described below.

- **Level 1 (Recall).** This level requires the recall of information such as a fact, definition, term, or simple procedure. These items require students only to demonstrate a rote response, use a well-known formula, or follow a set procedure.
- **Level 2 (Skill/Concept).** This level requires mental processing beyond that of recalling or reproducing a response. These items require students to make some decisions about how to approach the item.
- **Level 3 (Strategic Thinking).** This level requires reasoning, planning, and using evidence. These items require students to handle more complexity and abstraction than items at the previous two levels.

It is important that the Montana CRT in science measure a range of depths of knowledge. Table 3-13 shows the percent and point ranges of the three depth-of-knowledge levels used on the CRT in science.

Table 3-13. 2009–10 MontCAS:
Depth-of-Knowledge (DOK) Percent and Distribution by Level

DOK level	*Percent range*	*Point range*
1	20% to 26%	12 to 16 points
2	61% to 69%	37 to 42 points
3	5% to 20%	3 to 12 points

3.4.6 Use of Calculators and References Sheets

Calculators are not used or needed when taking the science tests. There are no science reference sheets.

3.5 TEST DEVELOPMENT PROCESS

3.5.1 Item Development

Items used on the Montana CRT are developed and customized specifically for use on the Montana CRT and are consistent with Montana content standards, benchmarks, and grade-level expectations. Measured Progress test developers work with Montana educators to verify the alignment of items to the appropriate Montana content standards.

The development process combined the expertise of Measured Progress test developers and committees of Montana educators to help ensure items meet the needs of the CRT program. All items used on the common portions of the Montana CRT program were reviewed by a committee of Montana content experts and by a committee of Montana bias experts. Tables 3-14 through 3-17 show the numbers of items developed within each content area for the 2009–2010 Montana CRT.

Table 3-14. 2009–10 MontCAS: Total Numbers of Items Developed by Content Area—Grades 3–8 and 10

Grade	*Mathematics*	*Reading*	*Science*
3	76	168	
4	76	168	78
5	76	168	
6	76	168	
7	76	168	
8	76	168	78
10	76	168	78

Table 3-15. 2009–10 MontCAS: Annual Reading Item Development—Grades 3–8 and 10

Passages	*MC*	*CR*
2 long literary passages	40	4
2 long informational passages	40	4
4 short literary passages	40	0
4 short informational passages	40	0
12 total passages	160	8

MC = multiple-choice; CR = constructed-response

Table 3-16. 2009–10 MontCAS: Annual Mathematics Item Development—Grades 3–8 and 10

MC	*SA*	*CR*
60	8	8

MC = multiple-choice; SA = short-answer; CR = constructed-response

Table 3-17. 2009–10 MontCAS: Annual Science Item Development—Grades 4, 8, and 10

MC	*CR*
75	3

MC = multiple-choice; SA = short-answer; CR = constructed-response

Table 3-18 provides an overview of the item development process for common and field-test items, including the administration of the operational tests.

Table 3-18. 2009–10 MontCAS: Item Development Process Overview

Development step	*Step details*
Select reading passages and conduct external review for bias and sensitivity issues (December 2007)	Measured Progress test developers located potential reading passages. Reading passages were reviewed for bias and sensitivity issues before the development of reading items.
Develop items (January through May 2008)	Measured Progress test developers developed reading and mathematics items.
Item review for content appropriateness and for bias and sensitivity issues (May 2008)	Committees of Montana educators reviewed reading, mathematics, and science field-test items.
Edit items (summer 2008)	Montana educators' recommended changes were incorporated into the items.
Field-test items (spring 2009)	Embedded field-test items were administered to a sample of students (minimum of 2,500 students per item) along with the 2009 operational test.
Item Selection Meeting (July 2009)	Measured Progress test developers and Montana educators selected common items for the spring 2010 operational CRT tests.
Operational test items (March 2010)	Items are now part of the common item set and are used to determine student scores. Another embedded field test was also administrated.

3.5.2 Item Reviews at Measured Progress

A test developer within each content area reviewed items for:

- item integrity, including content and structure, appropriateness to designated content area, format, clarity, possible ambiguity, and single correct answer.
- appropriateness and quality of reading selections and graphics.
- appropriateness of scoring guide descriptions and distinctions.
- whether the item is measuring the intended content standard.
- completeness of associated item documentation (e.g. scoring guide, content codes, key, grade level, depth of knowledge, and contract identified).
- appropriateness for the designated grade level.

3.5.3 Item Reviews at State Level

All passages and items were reviewed in Montana. In December 2007, the Montana Passage Review Committee met to review passages that would be developed for the 2009–10 CRT administration. The committee consisted of teachers and education specialists from across the state. In May 2008, Montana educators from across the state reviewed field-test items for content appropriateness, alignment to standards, depth of knowledge, and grade-level appropriateness.

3.5.4 Bias and Sensitivity Review

Bias review is an essential component of the development process. During the bias review process, reading passages and items from all content areas were reviewed by a committee of Montana educators. Items were examined for issues that might offend or dismay students, teachers, or parents. Including such groups in the development of assessment items and materials can avoid many unduly controversial issues, and unfounded concerns can be allayed before the test forms are produced.

3.5.5 Reviewing and Refining

Recommended changes from the Item Review and Bias and Sensitivity meetings were incorporated into the items by Measured Progress test developers.

3.5.6 Item Editing

Measured Progress editors then reviewed and edited the items to ensure adherence to sound testing principles and to style guidelines in the *Chicago Manual of Style, 15th ed.* These principles include the stipulations that items

- demonstrate correct grammar, punctuation, usage, and spelling;
- are written in a clear, concise style;
- contain unambiguous explanations that tell students what is required to attain a maximum score;
- are written at a reading level that allows students to demonstrate their knowledge of the subject matter being tested regardless of reading ability;
- exhibit high technical quality regarding psychometric characteristics;
- have appropriate answer options or score-point descriptors; and
- are free of potentially insensitive content.

3.5.7 Item Selection and Operational Test Assembly

In July 2009, Measured Progress test developers met with Montana educators to select common items. In preparation for the meeting, the test developers and psychometricians at Measured Progress considered the following in selecting sets of items to propose for the common item set to be used on the 2010 assessment:

- **Content coverage/match to test design and blueprints**. The test designs and blueprints stipulate a specific number of multiple-choice and constructed-response items for each content area. Item selection for the embedded field test was based on the number of items in the existing pool of items that are eligible for the common.

- **Item difficulty and complexity**. Item statistics drawn from the data analysis of previously field-tested items were used to ensure similar levels of difficulty and complexity from year to year as well as quality psychometric characteristics.
- **"Cueing" items**. Items were reviewed for any information that might "cue" or provide information that would help to answer another item.

At the meeting, the Montana educators reviewed the proposed sets of items and made the final selection of items for the common.

The test developers then sorted and laid out the items into test forms. During assembly of the test forms, the following criteria were considered:

- **Key patterns**. The sequence of keys (correct answers) was reviewed to ensure that their order appeared random.
- **Option balance**. Items were balanced across forms so that each form contained a roughly equivalent number of key options (*A*s, *B*s, *C*s, and *D*s).
- **Page fit**. Item placement was modified to ensure the best fit and arrangement of items on any given page.
- **Facing-page issues**. For multiple items associated with a single stimulus (reading passages) and multiple choice items with large graphics, consideration was given to whether those items needed to begin on a left- or right-hand page and to the nature and amount of material that needed to be placed on facing pages. These considerations served to minimize the amount of page flipping required of students.
- **Relationships among forms**. Although field-test items differ from form to form, these items must take up the same number of pages in all forms so that sessions begin on the same page in every form. Therefore, the number of pages needed for the longest form often determines the layout of each form.
- **Visual appeal**. The visual accessibility of each page of the form was always taken into consideration, including such aspects as the amount of "white space," the density of the test, and the number of graphics.

3.5.8 Operational Test Draft Review

After the forms were laid out as they would appear in the final test booklets, the forms were again thoroughly reviewed by Measured Progress editors to ensure that the items appeared exactly as intended. Any changes made during test construction were reviewed and approved by the test developer.

3.5.9 Alternative Presentations

Form 1 for the grade 3 through 8 and 10 tests was translated into Braille by National Braille Press, a subcontractor that specializes in test materials for blind and visually impaired students. In addition, Form 1 for each grade was adapted into a large-print version.

3.6 TEST SESSIONS

The Montana CRT was administered during the spring of 2010 during a four-week period from March 1, 2010 to March 24, 2010. Reading and mathematics tests were administered in grades 3 through 8 and 10, and science tests were administered in grades 4, 8, and 10. Schools were able to schedule testing sessions at any time during the four-week period, provided they followed the sequence detailed in the scheduling guidelines in the *Test Administrator's Manual.* Schools were asked to schedule makeup tests for students who were absent from initial test sessions during the testing window.

CHAPTER 4. TEST ADMINISTRATION

4.1 RESPONSIBILITY FOR ADMINISTRATION

As indicated in the *Test Coordinator's Manual*, principals and/or their designated school test coordinators are responsible for the proper administration of the CRT. This report was used to ensure the uniformity of administration procedures from school to school.

4.2 ADMINISTRATION PROCEDURES

School test coordinators were instructed to read the *Test Coordinator's Manual* prior to testing and to be familiar with the instructions given in the *Test Administrator's Manual*. The *Test Coordinator's Manual* provides each school with checklists to help prepare for testing. The checklists outline tasks to be performed before, during, and after test administration. In addition to providing these checklists, the *Test Coordinator's Manual* outlines the nature of the testing materials sent to each school, how to inventory the materials, how to track the materials during administration, and how to return the materials once testing was complete. The *Test Coordinator's Manual* also contains information about including or excluding students. The *Test Administrator's Manual* includes checklists for administrators to prepare themselves, their classrooms, and their students for administration of the test. The *Test Administrator's Manual* contains sections that detail the procedure to be followed for each test session, and it contains instructions for preparing the materials prior to giving them to school test coordinators for return to Measured Progress.

The Montana CRT is an untimed assessment; however, guidelines or ranges were provided in the *2009 Test Coordinator's Manual* and the *2009 Test Administrator's Manual* based on the following estimates of the time it takes an average student to respond to each type of item on the test:

- Multiple-choice items—1 minute per item
- Short-answer items—2 minutes per item
- Constructed-response items—10 minutes per item

The provided guidelines suggested scheduling 45–55 minutes per test session (50–60 minutes for grade 10 students). The guidelines also suggested scheduling a break between each of the three sessions in each content area to prevent fatigue.

While the guidelines for scheduling were based on the assumption that most students would complete the test within the estimated amounts of time, each test administrator was asked to allow additional time for students who needed additional time to complete the test. If additional classroom space was not available for this purpose, schools were encouraged to use another space, such as a guidance office. If other areas were not available, the guidelines recommended scheduling each classroom used for test administration for the maximum possible amount of time.

4.3 PARTICIPATION REQUIREMENTS AND DOCUMENTATION

All students were expected to participate in the CRT; however, the scores of students in the following categories were excluded from the calculation of averages:

- foreign exchange students
- students not enrolled in an accredited Montana school (for example, home-schooled students)
- students enrolled in a private accredited school
- students enrolled in a private non-accredited school
- students enrolled in a private non-accredited Title 1 school
- students enrolled part-time (less than 180 hours) taking a mathematics or reading course
- first year in U.S. LEP students were required to participate in the mathematics assessment only
- students who took the CRT using a "nonstandard" accommodation

A summary of this information is shown in Table 4-1, which was published in the *Test Administrator's Manual* and the *Test Coordinator's Manual*.

Table 4-1. 2009–10 MontCAS: Summary of Eligibility for Exclusion from the CRT

Excluded from averages	*MUST participate*	*MAY participate*
Foreign exchange students	Yes	
Students not enrolled in an accredited Montana school		Yes
Students enrolled in a private accredited school	Yes	
Students enrolled in a private non-accredited school		Yes
Students enrolled in a private non-accredited Title I school		Yes
Students enrolled part-time (less than 180 hours) taking a mathematics or reading course		Yes
Reading: first year in U.S. LEP students		Yes
Mathematics: first year in U.S. LEP students	Yes	

Staff members coded information about exclusion, if applicable; in the student response booklet after testing was completed. The *Test Coordinator's Manual* and *Test Administrator's Manual* provide detailed instructions for coding exclusions and accommodations. In addition, testing exclusions were discussed thoroughly in the pre-administration training audio CD (see Appendix A: Reporting Decision Rules).

Tables 4-2 through 4-4 below show a summary of participation on the 2009–10 Montana CRT by demographic category for each content area.

Table 4-2. 2009–10 MontCAS: Summary of Participation by Demographic Category – Mathematics

Description	*Number:*			*Percent tested*
	Enrolled	*Approved exemptions*	*Tested*	
Special Education	8665	715	7846	98.69
Title 1	26394	318	25929	99.44
Low Income	31573	796	30571	99.33
American Indian	9094	552	8459	99.03
Asian	719	34	684	99.85
Hispanic	2137	59	2061	99.18
Black or African American	896	24	867	99.43
White, Non-Hispanic	62733	1517	60929	99.53
Native Hawaiian/Other Pacific Islander	231	9	221	99.55
Female	36672	978	35521	99.52
Male	39158	1237	37700	99.42
Limited English Proficient	2082	108	1957	99.14
Migrant	133	0	133	100.00
Plan 504	522	8	503	97.86
All Students	76300	2666	73240	99.46

Table 4-3. 2009–10 MontCAS: Summary of Participation by Demographic Category – Reading

Description	*Number:*			*Percent tested*
	Enrolled	*Approved exemptions*	*Tested*	
Special Education	8603	714	7784	98.67
Title 1	26462	320	26000	99.46
Low Income	31507	794	30511	99.34
American Indian	9073	552	8437	99.01
Asian	721	34	684	99.56
Hispanic	2131	60	2054	99.18
Black or African American	891	24	862	99.42
White, Non-Hispanic	62674	1515	60885	99.55
Native Hawaiian/Other Pacific Islander	230	9	220	99.55
Female	36665	980	35511	99.51
Male	39075	1234	37631	99.45
Limited English Proficient	2074	108	1946	98.98
Migrant	133	0	133	100
Plan 504	522	8	504	98.05
All Students	76214	2665	73165	99.48

Table 4-4. 2009–10 MontCAS: Summary of Participation by Demographic Category – Science

Description	*Number:*			*Percent tested*
	Enrolled	*Approved exemptions*	*Tested*	
Special Education	3765	326	3374	98.11
Title 1	131	0	128	97.71
Low Income	12786	353	12287	98.83
American Indian	3718	223	3436	98.31
Asian	334	20	313	99.68

continued

Description	Number:			Percent tested
	Enrolled	Approved exemptions	Tested	
Hispanic	901	29	861	98.74
Black or African American	352	15	334	99.11
White, Non-Hispanic	27300	771	26324	99.23
Native Hawaiian/Other Pacific Islander	110	4	105	99.06
Female	15814	481	15204	99.16
Male	16901	581	16169	99.07
Limited English Proficient	833	40	778	98.11
Migrant	59	0	59	100
Plan 504	273	7	258	96.99
All Students	32910	1242	31388	99.12

4.3.1 Students with Disabilities

All students with special needs participate in the CRT assessment program, either by taking the regular CRT or CRT-Alternate Assessment if they meet the eligibility criteria.

Form 1 for the grade 3 through 8 and 10 tests was enlarged to 18 point font for visually impaired students as well as translated into Braille by National Braille Press, a subcontractor that specializes in test materials for blind students. Students with special needs and LEP students are often given test accommodations.

4.4 ADMINISTRATOR TRAINING

The OPI hosted a test-administration workshop in Helena, Montana, on January 28–29, 2010. The workshop was well attended, but attendance of system and school test coordinators was not mandatory. OPI and Measured Progress staff members hosted six sessions that covered test accommodations, student information system (AIM) updates, CRT materials and administration, CRT-Alternate materials and administration, online reporting, and test security. Each session was presented six times so that participants could be educated on all facets of test administration.

In addition to the workshop and the distribution of the *2010 Test Coordinator's Manual* and *Test Administrator's Manual*, the OPI and Measured Progress produced and distributed one audio PowerPoint presentation, "Spring 2010: CRT and CRT-ALT Overview and Update of System and School Test Coordinators," to each system and school test coordinator. Training materials and the audio PowerPoint presentation were also posted on the OPI's Web site: http://www.opi.mt.gov. The training CD provided system and school test coordinators who were unable to attend the administration workshops with the training materials. The CDs also served as useful tools for training both system and school personnel.

4.5 DOCUMENTATION OF ACCOMMODATIONS

The 2010 CRT Accommodations Manual and the accommodations training PowerPoint, *Guidelines and Procedures for CRT Accommodations*, were produced by the OPI and were included on the CRT training CD provided to each system and school in the first shipment received by systems in early February 2010 from Measured Progress. General instructions regarding accommodation usage and a list of available accommodations were included in the *2010 Test Coordinators Manual*.

Standard accommodations were available to all students on the basis of individual needs and regardless of disability status. Decisions regarding standard accommodations were made by the student's educational team on an individual basis, consistent with either previous accommodation decisions for the student or current educational needs. Accommodations usage was to be consistent with those used during the student's regular classroom instruction and assessment for at least three months prior to testing.

Nonstandard accommodations were offered to students with disabilities only if the accommodations are specified in the student's IEP. If a student was assessed with a nonstandard accommodation, the student was considered a nonparticipant when calculating the participation rate for AYP purposes. In addition to the student being considered a nonparticipant, the student's score from the assessment is not included in calculating the proficiency rate for AYP.

Table 4-5 below shows the number of students at each subject and grade who were tested with and without accommodations. In addition, frequencies of accommodations used by accommodation type are presented in Appendix B.

Table 4-5. Numbers of Students Tested With and Without Accommodations by Subject and Grade

Subject	*Grade*	*Number of students tested:*	
		With accommodations	*Without accommodations*
Mathematics	3	1847	8698
	4	1875	8388
	5	1675	8839
	6	1439	9003
	7	1176	9281
	8	1087	9542
	10	741	9649
Reading	3	1802	8707
	4	1782	8444
	5	1619	8872
	6	1426	8993
	7	1163	9297
	8	1094	9543
	10	759	9664
Science	4	1569	8697
	8	1041	9630
	10	780	9671

4.6 TEST SECURITY AND ADMINISTRATION IRREGULARITIES

Test coordinators and administrators are prohibited from disclosing the contents of CRT assessments. Under no circumstances should test booklets or marked answer booklets be circulated among faculty, administrators, or other persons.

All system test coordinators and school principals received the *OPI Guidelines and Procedures for Test Security*. This OPI publication was made available to system superintendents, principals, and test administrators for the purpose of outlining the reporting procedures for security and administration violations. All concerns about breaches of test security or noncompliance with test administration procedures were to be reported immediately to the principal, system test coordinator, and state assessment director.

4.7 TEST ADMINISTRATION WINDOW

The Montana CRT was administered during the spring of 2010 during a four-week period from March 1, 2010 to March 24, 2010. Reading and mathematics tests were administered in grades 3 through 8 and 10, and science tests were administered in grades 4, 8, and 10. Schools were able to schedule testing sessions at any time during the four-week period, provided they followed the sequence detailed in the scheduling guidelines in the *Test Administrator's Manual.* Schools were asked to schedule makeup tests for students who were absent from initial test sessions during the testing window.

4.8 SERVICE CENTER

To address testing concerns, Measured Progress established a help desk dedicated to the Montana CRT. Service Center support is an essential element to the successful administration of large-scale assessments. It provides a central location that individuals in the field can call via a toll-free number to request assistance, report problems, or ask specific questions.

The Measured Progress help desk provided support during all phases of the testing window. It was staffed at varying levels, based on need and the volume of calls received, from 8:00 A.M. to 4:00 P.M. MST. At a minimum, the help desk consisted of a product support specialist responsible for receiving, responding to, and tracking calls and e-mails, and routing issues to the appropriate person(s) for resolution. In addition, the program manager and/or program assistant addressed communications that required a higher level of program support.

During the period between February 19, 2010, when the testing materials were delivered to schools, and April 6, 2010 when the materials were returned to Measured Progress, the Service Center received 130 calls. The majority of these calls were to order additional materials for students who enrolled after materials were shipped and to arrange for UPS to pick up the materials after testing. The service center staff also responded to administration questions and referred policy questions regarding test security or accommodations usage to the OPI.

CHAPTER 5. SCORING

Accurate and timely scoring of constructed-response short-answer, and multiple-choice items is an important process in any successful assessment program. This chapter defines the scope and processes of Measured Progress's Scoring Services for the 2009–2010 Montana CRT.

5.1 MACHINE-SCORED ITEMS

Preceding the arrival of the Montana CRT student response booklets, Measured Progress prepared customized scanning programs to enable selective reading of all scannable materials including student identification and demographics and to electronically format the scanned information.

Once the student answer documents were received from each Montana school following test administration, Measured Progress optically scanned each page from every student booklet to create digital images of the entire document. Every page was bar coded so that the scores applied to each item could be linked to the correct student, school, and district. Student responses were then imported into iScore™, Measured Progress's proprietary, image-based scoring system, for secure processing and scoring. By using this image scoring system, Measured Progress was able to increase reliability and productivity as well as monitor and maintain quality control.

Student multiple-choice response data was machine-scored at the same time that student constructed-response and short-answer items were scanned into iScore for person-scoring. Multiple-choice items were compared to scoring keys via item analysis software. Correct multiple-choice answers were assigned a score of one point and incorrect answers were assigned zero points. Student multiple-choice responses consisting of multiple marks and blank responses were also assigned zero points. The total number of responses scanned and scored is presented in Table 5-1 for each assessed grade in each content area.

Student responses that could not be physically scanned (e.g., documents damaged during administration or shipment), were physically reviewed and scored on an individual basis by trained, qualified staff. These scores were linked to the student's demographic data and merged with the student's scoring file by Measured Progress's Data Processing department.

Table 5-1. 2009–10 Montana CRT Number of Responses Scanned and Scored

Content area	*Grade*	*Number of responses scanned and scored*
Mathematics	3	54,060
	4	52,730
	5	53,730
	6	52,539
	7	53,560
	8	54,586
	10	54,460
Reading	3	21,952
	4	21,338
	5	21,738

continued

Content area	Grade	Number of responses scanned and scored
Reading	6	21,344
	7	21,669
	8	22,044
	10	21,910
Science	4	21,338
	8	22,044
	10	21,910

5.2 PERSON-SCORED ITEMS

Scanned images of open-response items were processed and organized into item-specific groups in preparation for person-scoring by iScore. iScore's secure, Web-based application provided qualified staff, including readers and their leadership staff, password-protected access for reading and scoring electronic student responses at one or multiple scoring sites without compromising confidentiality. The digital image clip information of constructed-response and short-answer responses allowed iScore to replicate student responses just as they appeared on the originals and to display the replicated responses on individual monitors for person-scoring. In addition, the processes of item benchmarking, reader training, scoring, editing/cleanup, and reporting were all accomplished electronically and without further reference to the originals.

Organized by iScore in this way, qualified readers were able to view only one response from a single item at a time. Because item responses were tracked and distributed among groups of readers by iScore, each response in an individual student's response booklet was able to be assigned to and be scored by a different reader. This maximization of the number of readers per student response booklet effectively minimized bias errors due to reader sampling.

Leadership staff, on the other hand, had constant, albeit view-only, access to all of the imaged responses from a student's booklet for whenever necessary. The actual test booklets and answer documents were also available to the content chief reader and the iScore operational manager (see section on "Scoring Location and Staff").

To ensure the security of constructed-response and short-answer items and responses scored, all scoring activities in iScore were performed "blind": i.e., without student names, district, and/or school information visible or able to be associated with responses or raw scores. During scoring, iScore distributed images of student responses to the computer monitors of readers located at one of Measured Progress's scoring facilities. When iScore sent an image of a student response to an individual reader's computer monitor, the reader evaluated the response and recorded the score via keypad or mouse entry. Once the score was entered, a new response appeared immediately on the screen.

Although iScore is based on conventional scoring techniques, it also offered the following benefits;

- provided leadership staff with real-time information about group and individual level performance including scoring accuracy and consistency as well as overall process monitoring and reporting,
- ensured the randomized distribution of student responses among readers during scoring and automatically assigned student responses to one or more scorers for interrater agreement monitoring,
- permitted password-only access limited to those solely in the employ of Measured Progress working within a qualified scoring or scoring management capacity,
- maintained student anonymity and confidentiality by masking student biographical information from viewers,
- offered immediate access to samples of student responses and scores for reporting and analysis,
- offered early access to subsets of data for tasks such as standard setting, and
- reduced material handling, which saved time and labor while enhancing the security of materials.

The iScore database, its control operation, and its administrative offices were all based in Dover, New Hampshire. The iScore system monitored accuracy, reliability, and consistency across all Measured Progress scoring facilities. To ensure that scoring information and updates were equally shared and implemented across all scoring facilities, constant communication and coordination was accomplished daily via e-mail, telephone, fax, and secure Web-based networks.

5.2.1 Scoring Location and Staff

Scoring Location

Scoring the 2009–10 Montana CRT program took place in Dover, New Hampshire. The overview of scoring operations is presented by content and grade in Table 5-2.

Table 5-2. 2009–10 Montana CRT Scoring Locations by Content and Grade

Content/grade level	*Dover, NH (day shift)*	*Dover, NH (night shift)*
Mathematics Grade 3		X
Mathematics Grade 4		X
Mathematics Grade 5	X	
Mathematics Grade 6	X	
Mathematics Grade 7	X	
Mathematics Grade 8	X	
Mathematics Grade 10	X	
Reading Grade 3		X
Reading Grade 4		X
Reading Grade 5	X	
Reading Grade 6	X	
Reading Grade 7	X	

continued

Content/grade level	*Dover, NH (day shift)*	*Dover, NH (night shift)*
Reading Grade 8		X
Reading Grade 10		X
Science Grade 5		X
Science Grade 8		X
Science Grade 10		X

Scoring Staff

Staffing for the 2009–10 Montana CRT implemented low scoring-leadership-to-reader ratios and was carried out by the following Measured Progress staff members:

- Scoring project manager, who oversaw the overall contract from a scoring perspective and acted as a liaison with contract management staff, data analysis staff, and the client while managing the content experts (chief readers, quality assurance coordinators, etc.).
- Chief readers, who prepared benchmarking/training materials and led the review and client approval of materials, working closely with Curriculum & Assessment specialists and Montana educators. Chief readers trained, qualified, and monitored readers during the scoring process; supervised quality assurance coordinators, senior readers, and readers; and monitored scoring accuracy and consistency. The ratio of chief readers to the scoring project manager was 3:1.
- Quality assurance coordinators (QACs), who managed the training and benchmarking of grades and items within the Montana CRT. QACs trained, qualified, and monitored readers during the scoring process; supervised senior readers and readers; and monitored scoring accuracy and consistency. The ratio of QACs to chief readers was 7:1.
- Senior readers (SRs), who supervised readers during the scoring process, monitoring scoring accuracy and consistency while managing quality control measures via iScore. The ratio of SRs to QACs was 1:1.
- Readers, who were qualified, temporary staff members performing the bulk of scoring work, evaluating and scoring student responses according to the Montana CRT guidelines provided for each grade level and content area scored. Readers received the same orientation and training as direct hires. The ratio of readers to SRs was 9:1.

5.2.2 Reader Recruitment and Qualifications

In preparation for scoring the 2009–10 Montana CRT, Measured Progress actively sought and recruited readers to represent a diverse spectrum of educational, professional, and ethnic populations. The customary cross-section of readers employed included content specialists such as editors, business professionals, scientists, authors, graduate school students, and both current and retired educators.

Although the employment of readers holding a four-year college degree or higher was preferred, all readers were required to have successfully completed a minimum of at least two years of college and to have demonstrated knowledge of the content area they scored. All readers were required to submit documentation (i.e., college transcript and/or resume) of their qualifications.

For training and qualification, readers were placed at grade levels and in content areas that matched their areas of experience and expertise. Reader demographic information (gender, educational, and ethnic background, etc.) was electronically documented for reporting. All readers were subject to stringent nondisclosure requirements and supervision and were required to sign a nondisclosure/confidentiality agreement. Table 5-3 summarizes the educational credentials of the 2009–10 Montana CRT readers and QACs.

Table 5-3. 2009–10 Montana CRT Educational Credentials of Readers and QACs

Readers				
Description	*Dover, NH (1st shift)*	*Dover, NH (2nd shift)*	*Total*	*Percent*
Less then 48 college credits	0	0	0	0.0%
48+ college credits	5	4	9	4.6%
Associate's degree	4	7	11	5.6%
Bachelor's degree	76	44	120	60.5%
Master's degree	32	21	53	26.8%
Doctorate	4	1	5	2.5%
Total	121	77	198	100.0%

QACs				
Description	*Dover, NH (1st shift)*	*Dover, NH (2nd shift)*	*Total*	*Percent*
Less then 48 college credits	0	0	0	0.00%
48+ college credits	0	0	0	0.00%
Associate's degree	0	0	0	0.00%
Bachelor's degree	12	9	21	58.3%
Master's degree	10	4	14	38.9%
Doctorate	1	0	1	2.8%
Total	23	13	36	100.0%

5.2.3 Methodology for Scoring Polytomous Items

Possible Score Points

The ranges of possible score points for the different Polytomous items found on the 2009–10 Montana CRT are shown in Tables 5-4 and 5-5.

Table 5-4. 2009–10 Montana CRT Short-Answer Item Scoring Guide

Score point	*Description*
1	The student's response provides a complete and correct answer.
0	The student's response is totally incorrect or too minimal to evaluate.
B	Blank/no response.

Table 5-5. 2009–10 Montana CRT Constructed-Response Item Scoring Guide

Score point	*Description*
4	The student completes all important components of the task and communicates ideas clearly.
	The student demonstrates in-depth understanding of the relevant concepts and/or processes.
	When instructed to do so, the student chooses more efficient and/or sophisticated processes.
	When instructed to do so, the student offers insightful interpretations or extensions (e.g., generalizations, applications, and analogies).
3	The student completes the most important components of the task and communicates clearly.
	The student demonstrates understanding of major concepts even though he/she overlooks or misunderstands some less important ideas or details.
2	The student completes most important components of the task and communicates those clearly.
	The student demonstrates that there are gaps in his/her conceptual understanding.
1	The student shows minimal understanding.
	The student addresses only a small portion of the required task(s).
0	The student's response is totally incorrect or irrelevant.
B	Blank/no response.

Condition Codes

When numerical score-point parameters did not apply to a student response, readers had the option of designating one the following options**:**

- Blank response (empty entry without an attempt at responding to the question)
- Unreadable response (response is too illegible or faint to accurately interpret)
- Wrong Location (a relevant response entered into the space reserved for a different item)

- Non-English response (a response written entirely in a language other than English)

Unreadable and Wrong Location responses were resolved by consulting the original test booklet and/or by identifying the correct location.

5.2.4 Reader Training

For each item scored on the 2009–10 Montana CRT, Measured Progress readers were required to demonstrate their scoring ability by participating in training sessions specific to each student response item scheduled to be scored. The scoring project began with an introduction of the onsite scoring staff and providing an overview of the Montana CRT program's purpose and goals (including discussion about document security, student confidentiality, the proprietary nature of testing materials, scoring materials, and iScore procedures).

Actual training began with groups of readers organized into content-, grade-, and item-specific group assignments. Each reader per group was provided a personal hard copy of item-specific training materials distributed at the beginning of each work session and accounted for secure collection at the end of each work session. During training, readers were strongly encouraged to take notes and highlight their own hard copies of the training materials.

For each item trained, the QAC assigned to the item commenced reader training by reviewing and discussing the item-specific scoring guide. The training QAC demonstrated the process of applying the item's scoring guide and score point descriptors to the exemplars found in the subsequent Anchor and Training Sets before attempting to demonstrate scoring accuracy in the Qualifying Set.

Anchor Set

This is a set of responses approved by the respective content specialists for reading, mathematics, or science representing the Montana Office of Public Instruction. Each Anchor Set contained one OPI-approved sample response per score point considered to be a mid-range exemplar. This set occasionally included a second sample response if there was more than one plausible way to illustrate the merits and intent of a score point. Responses in the Anchor Sets were typical, rather than unusual or uncommon; solid, rather than controversial or borderline; and true, meaning that their scores could not be changed except by the OPI and Measured Progress test development specialists.

Each Anchor Set response was read aloud to readers by the training QAC. Training QACs facilitated group discussion of responses in relation to the scoring guide and score point descriptors to help readers internalize the typical characteristics of score points. The Anchor Set served as a reference for readers as they went on to scoring sample responses in the Training Set that followed.

Training Set

Next, readers practiced applying the scoring guide and Anchor Set to responses in the Training Set. The Training Set typically included 6 to 10 student responses designed to help establish both the full score point range and the variation of possible responses within each score point. The Training Set often included unusual responses that were less clear or solid (e.g., briefer than normal, employing atypical approaches, simultaneously containing very low and very high attributes, and written in ways difficult to decipher).

Responses in the Training Set were presented to readers without scores and in a randomized score point order. Once readers had independently read and applied their score to a Training Set response, the training QAC would poll readers and then announce the actual response score. The QAC then responded to reader questions and/or comments while pointing out particular scoring issues at hand (e.g., the borderline between two score points). Throughout each item training, the QAC continually routed reader discussion of score points back to the Anchor Set and scoring guide. After the Training Set had been completed, readers were required to demonstrate scoring accuracy using Qualifying Sets assembled for constructed-response items.

Qualifying Set

Following participating in each item training session, readers were administered a Qualification Set of Committee-Reviewed Responses (CRRs) assigned to each item in the reader's content area. Each Qualifying Set was composed of ten, preselected, previously scored responses chosen as clear illustrations of score point examples that would measure which readers had adequately internalized item training before being able to score live student responses. These CRRs were selected by scoring leadership and randomly distributed to each reader via iScore during qualification.

In order to qualify on a Qualification Set, readers were required to demonstrate a scoring accuracy level of at least eighty percent (80%) exact agreement (i.e., exactly match scores on at least 8 of the 10 CRRs) and at least ninety percent (90%) exact or adjacent agreement (i.e., exactly match or be within one score point on 9 or 10 of the 10 CRRs). In other words, readers were allowed 1 discrepant score (i.e., 1 score out of the 10 CRRs that was more than one score point from the CRR score) provided they had at least 8 exact scores.

Once a group of readers successfully qualified on a particular item, responses for that item in iScore were randomly assigned and presented to them on their computer monitors, one response at a time. Readers unable to qualify on the first Qualification Set received QAC retraining (see section on "Retraining") and a subsequent opportunity at qualification on a second Qualification Set. Readers unable to qualify on the second Qualification Set were not eligible to score that item.

(Note: In order to be eligible to score SA1 mathematics items in grades 3 through 8 and 10, readers were required to qualify on at least one mathematics constructed-response item for that grade.)

Retraining

Readers unable to qualify on the first Qualification Set received QAC retraining by reviewing their performance in relation to the item training materials. The QAC responded to reader questions and routed discussion of score points back to the Anchor Set and scoring guide. Readers were then allowed the opportunity at qualification on a second Qualification Set. Readers unable to qualify on the second Qualification Set were not eligible to score that item. Table 5-6 depicts the accuracy and qualification percentages of the reader applicants.

Table 5-6. 2009–10 Montana CRT Scoring Accuracy and Qualification Statistics

Content area	*Grade*	*Item*	*Average percent exact agreement for embedded CR sets*	*Average percent exact agreement for double-blind scoring*	*Number of readers taking qualification sets*	*Number successfully qualifying*	*Percent successfully qualifying*
Mathematics	3	23	NA	96.0	NA	NA	NA
		24	NA	98.1	NA	NA	NA
		25	97.0	91.5	16	16	100.0
		48	NA	96.4	NA	NA	NA
		72	93.8	90.7	16	16	100.0
	4	23	NA	98.2	NA	NA	NA
		24	NA	99.7	NA	NA	NA
		25	92.5	92.6	20	20	100.0
		48	NA	98.8	NA	NA	NA
		72	88.8	90.3	21	21	100.0
	5	23	NA	97.1	NA	NA	NA
		24	NA	96.1	NA	NA	NA
		25	86.7	87.4	11	11	100.0
		48	NA	98.5	NA	NA	NA
		72	89.6	90.9	11	11	100.0
	6	18	NA	94.8	NA	NA	NA
		19	NA	93.8	NA	NA	NA
		20	NA	93.9	NA	NA	NA
		23	78.5	91.5	13	13	100.0
		73	87.6	83.0	8	8	100.0
	7	18	NA	97.1	NA	NA	NA
		19	NA	96.6	NA	NA	NA
		20	NA	96.6	NA	NA	NA
		23	94.6	93.9	13	10	76.9
		73	85.3	90.3	14	13	92.9
	8	18	NA	98.6	NA	NA	NA
		19	NA	97.6	NA	NA	NA
		20	NA	96.5	NA	NA	NA
		23	84.5	90.3	12	12	100.0

continued

Content area	Grade	Item	Average percent exact agreement for embedded CR sets	Average percent exact agreement for double-blind scoring	Number of readers taking qualification sets	Number successfully qualifying	Percent successfully qualifying
Mathematics	8	73[1]	-	-	-	-	-
	10	18	NA	98.1	NA	NA	NA
		19	NA	95.6	NA	NA	NA
		20	NA	98.4	NA	NA	NA
		23	72.4	91.5	31	25	80.6
		73	81.7	91.4	8	8	100.0
Reading	3	27	74.9	77.5	24	23	95.8
		81	71.3	81.3	21	20	95.2
	4	27	71.7	74.1	18	16	88.9
		81	74.8	74.5	20	19	95.0
	5	27	75.2	70.3	13	13	100.0
		81	84.1	75.5	13	13	100.0
	6	27	74.5	75.3	13	13	100.0
		81	72.1	79.2	13	13	100.0
	7	27	73.7	70.1	11	11	100.0
		81	82.7	74.6	14	11	78.6
	8	27	74.3	63.2	15	14	93.3
		81	81.8	70.7	13	13	100.0
	10	27	80.9	79.4	15	14	93.3
		81	88.7	84.8	18	18	100.0
Science	4	27	97.1	88.3	9	9	100.0
		81	98.2	88.6	9	9	100.0
	8	27	98.0	81.1	17	17	100.0
		81	90.5	82.9	23	23	100.0
	10	27	93.5	86.7	12	12	100.0
		81	85.0	92.0	15	14	93.3

5.2.5 Leadership Training

A core group of scoring leadership staff including QACs and SRs guided and monitored readers throughout the scoring process. Because quality control by QACs and SRs moderated the scoring process and maintained the integrity of scores, the individuals chosen to fill those positions were selected for their accuracy and consistency. In addition, the training QACs assigned to train readers were also selected for their ability to instruct and their content area specialization.

The purpose of leadership training was to ensure that QACs provided thorough and consistent training and feedback to readers. Chief readers were able to discuss item details and score point rationale within training materials in order to prepare scoring leadership for reader questions. Chief readers trained and

[1] Was not scored.

reviewed items with QACs, who in turn trained and reviewed items with their SRs. During actual item scoring, QACs trained and supervised readers and monitored SR accuracy and consistency. The SRs in turn supervised their own group of readers and monitored reader accuracy and consistency. Similar to readers, scoring leadership who performed quality control measures in iScore were also required to qualify using Qualification Sets by demonstrating a scoring accuracy level of at least eighty percent (80%) exact agreement and at least ninety percent (90%) exact or adjacent agreement.

5.2.6 Monitoring of Scoring Quality Control

iScore was pre-programmed to monitor individual reader accuracy and scoring consistency among readers on a constant basis. iScore's use of multiple monitoring techniques was critical to the process of live scoring, allowing readers who met or exceeded accuracy standards access to continue scoring. Reader accuracy and consistency was measured in iScore throughout the scoring process using the following methods and tools:

- Embedded Committee-Reviewed Responses (CRRs)
- Read-Behind Scoring
- Double-Blind Scoring
- Reader Arbitration

Embedded Committee-Reviewed Responses

Embedded Committee-Reviewed Responses (Embedded CRRs) are preselected, previously scored responses used to ensure that readers had adequately internalized item training and remained calibrated to the scoring standard during live scoring. Previous to scoring, scoring leadership selected Embedded CRRs for each item and loaded the examples into iScore ("embedded"). Each example represented images of actual student work and appeared no different than live student responses. During the first day of live scoring of each item, iScore randomly distributed thirty (30) Embedded CRRs to each reader. Embedded CRRs were employed for all constructed-response items and enabled scoring leadership to monitor reader accuracy and consistency as gauged by the known scores of the Embedded CRRs.

Readers with a disproportionate number of adjacent and/or discrepant scores in Embedded CRRs were subject to increased monitoring, additional read-behinds, consultation by scoring leadership, and/or retraining by the QAC. Following these measures, it was at the discretion of scoring leadership whether or when the reader could resume scoring. If the individual was allowed to resume scoring, scoring leadership carefully monitored these readers by increasing the number of read-behinds.

Read-Behind Procedures

Read-behind scoring refers to scoring leadership (typically a SR) scoring a response that was recently scored by a reader. The gain was an immediate, real-time "snapshot" of each reader's accuracy and consistency during scoring. SRs were required to perform read-behinds on each reader throughout each day and at any point during scoring. This practice was applied to all open-ended item types and performed on all readers as required.

Once called up in iScore by the SR, read-behind responses were selected by iScore and placed into the SR's read-behind queue. Readers were aware neither of iScore's selection nor which of their scored responses was to be reviewed by their SR. Likewise, SRs were not aware of the reader's score when iScore presented each read-behind response for their own review and eventual score. The SRs then applied their own score to the response before the reader's score was made viewable in iScore. This SR review and comparison of the two scores created the score of record determination (i.e., the reported score) as follows:

- If the reader and SR applied the same score (exact agreement), no action was necessary; the reader's score became the score of record.
- If the reader and SR scores differed by 1 point (adjacent), the SR's score became the score of record, thereby overriding the reader's score.
- If the reader and SR scores differed by more than 1 point (discrepant), the SR's score became the score of record, thereby overriding the reader's score.
- Readers with a disproportionate number of adjacent and/or discrepant scores in read-behinds were subject to increased monitoring, additional read-behinds, consultation by scoring leadership, and/or retraining by the QAC. Following these measures, it was at the discretion of scoring leadership whether or when the reader could resume scoring. If the individual was allowed to resume scoring, scoring leadership carefully monitored these readers by increasing the number of read-behinds. Table 5-7 outlines the resolution of reader scores using the read-behind procedure.

Table 5-7. 2009–10 Montana CRT Examples of Read-Behind Scoring Resolution

Reader score	*QAC/SR score*	*Score of record*
4	4	4
4	3	3*
4	2	2*

* QAC/SR's score.

Double-Blind Scoring

Scoring procedures for both constructed-response and short-answer item types included both single-scoring and double-scoring. Single-scored responses were scored by one reader. Double-scored responses

were scored "blindly" by two different readers, unaware of the other's score. These double-blind scores were monitored for interrater agreement accuracy and scoring consistency. A default minimum setting of two percent (2%) from all open-ended item types was double-blind scored. In addition, responses marked blank, unreadable, and non-English were automatically routed for double-blind scoring. Table 5-8 indicates the incidence frequency for which open-ended item responses from each content area were double-blind scored.

Table 5-8. 2009–10 Montana CRT Frequency of Double-Blind Scoring by Grade and Content

Grade	*Content*	*Responses double-blind scored*
3–8, 10	Mathematics	2% minimum
3–8, 10	Reading	2% minimum
4, 8, 10	Science	2% minimum
All	Unreadable responses	100%
All	Blank responses	100%
All	Non-English	100%

Reader Arbitration

When double-blind scores applied by two readers on a single response differed by more than 1 point (discrepancy), iScore placed the response into an arbitration queue for review and rescoring by the SR. Readers were aware neither of the discrepancy arbitration nor which of their scored responses was to be arbitrated. Likewise, the SR was not aware of either readers' scores when iScore presented the response for review. It was only after the SR had applied their own score to the response that the readers' scores were then made viewable. This SR review and rescoring of the response became the score of record, thereby overriding the readers' scores.

Readers with a disproportionate number of adjacent and/or discrepant scores in double-blind scoring were subject to increased monitoring, additional read-behinds, consultation by scoring leadership, and/or retraining by the QAC. Following these measures, it was at the discretion of scoring leadership whether or when the reader could resume scoring. If the individual was allowed to resume scoring, scoring leadership carefully monitored these readers by increasing the number of read-behinds. Table 5-9 displays the final summary statistics for double-blind scoring.

Table 5-9. 2009–10 Montana CRT Double-Blind Summary Statistics

Content area	*Grade*	*Number of responses scored*	*Total number of responses double-blind scored*	*Percentage of responses double-blind scored*
Mathematics	3	54,060	1,767	3.3
	4	52,730	1,737	3.3
	5	53,730	1,672	3.1
	6	52,539	2,378	4.5
	7	53,560	2,563	4.8
	8	54,586	1,795	3.3

continued

Content area	*Grade*	*Number of responses scored*	*Total number of responses double-blind scored*	*Percentage of responses double-blind scored*
Mathematics	10	54,460	3,953	7.3
Reading	3	21,952	813	3.7
	4	21,338	698	3.3
	5	21,738	640	2.9
	6	21,344	619	2.9
	7	21,669	826	3.8
	8	22,044	662	3.0
	10	21,910	1,109	5.1
Science	4	21,338	714	3.3
	8	22,044	1,007	4.6
	10	21,910	1,879	8.6

In the case that the individual was not allowed to resume scoring however, the content chief reader had the right to remove ("void") all of that reader's scores applied to the item from that day's work totals. Voided responses in iScore were returned back to the response queue and rescored by readers able to maintain the scoring accuracy standard. Table 5-10 summarizes the statistics relevant to voided or blocked readers.

Table 5-10. 2009–10 Montana CRT Voided or Blocked Reader Statistics

Content area	*Grade*	*Item*	*Number of readers with voided scores*	*Number of readers NOT allowed to continue scoring based upon other quality monitoring (read-behinds and double blinds)*
Mathematics	3	23	0	0
		24	0	0
		25	0	0
		48	0	0
		72	0	0
	4	23	0	0
		24	0	0
		25	0	0
		48	0	0
		72	0	0
	5	23	0	0
		24	0	0
		25	0	0
		48	0	0
		72	0	0
	6	18	0	0
		19	0	0
		20	0	0
		23	0	0
		73	1	0

continued

Content area	*Grade*	*Item*	*Number of readers with voided scores*	*Number of readers NOT allowed to continue scoring based upon other quality monitoring (read-behinds and double blinds)*
Mathematics	7	18	0	0
		19	0	0
		20	0	0
		23	0	0
		73	0	0
	8	18	0	0
		19	0	0
		20	0	0
		23	8	0
		73	0	0
	10	18	0	0
		19	0	0
		20	0	0
		23	0	0
		73	0	0
Reading	3	27	0	0
		81	0	0
	4	27	0	0
		81	0	0
	5	27	0	0
		81	0	0
	6	27	0	0
		81	0	0
Reading	7	27	0	0
		81	0	0
	8	27	0	0
		81	0	0
	10	27	0	0
		81	0	0
Science	4	23	0	0
		69	0	0
	8	23	0	0
		69	0	0
	10	23	3	0
		69	3	0

CHAPTER 6. CLASSICAL ITEM ANALYSIS

As noted in Brown (1983), "A test is only as good as the items it contains." A complete evaluation of a test's quality must include an evaluation of each item. Both *Standards for Educational and Psychological Testing* (AERA et al., 1999) and *Code of Fair Testing Practices in Education* (2004) include standards for identifying quality items. Items should assess only knowledge or skills that are identified as part of the domain being tested and should avoid assessing irrelevant factors. Items should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. In addition, items must not unfairly disadvantage students in particular racial, ethnic, or gender groups.

Both qualitative and quantitative analyses are conducted to ensure that Montana CRT items meet these standards. Qualitative analyses are described in earlier chapters of this report; this chapter focuses on quantitative evaluations. Statistical evaluations are presented in four parts: 1) difficulty indices, 2) item-test correlations, 3) differential item functioning (DIF) statistics, and 4) dimensionality analyses. The item analyses presented here are based on the statewide administration of the Montana CRT in spring 2010. Note that the information presented in this chapter is based on the items common to all forms, since those are the items on which student scores are calculated. (Item analyses are also performed for field-test items, and the statistics are then used during the item review process and form assembly for future administrations.)

6.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES

All multiple-choice, constructed-response, and short-answer items are evaluated in terms of item difficulty according to standard classical test theory practices. Difficulty is defined as the average proportion of points achieved on an item and is measured by obtaining the average score on an item and dividing it by the maximum possible score for the item. Multiple-choice and short-answer items are scored dichotomously (correct vs. incorrect) so, for these items, the difficulty index is simply the proportion of students who correctly answered the item. Constructed-response items are scored polytomously, meaning that a student can achieve a score of 0, 1, 2, 3, or 4. By computing the difficulty index as the average proportion of points achieved, the indices for the different item types are placed on a similar scale, ranging from 0.0 to 1.0 regardless of the item type. Although this index is traditionally described as a measure of difficulty, it is properly interpreted as an *easiness* index, because larger values indicate easier items. An index of 0.0 indicates that all students received no credit for the item, and an index of 1.0 indicates that all students received full credit for the item.

Items that are answered correctly by almost all students provide little information about differences in student abilities, but they do indicate knowledge or skills that have been mastered by most students. Similarly, items that are correctly answered by very few students provide little information about differences in student abilities, but may indicate knowledge or skills that have not yet been mastered by most students. In general, to provide the best measurement, difficulty indices should range from near-chance performance (0.25 for four-

option multiple-choice items or essentially zero for constructed-response or short-answer items) to 0.90, with the majority of items generally falling between around 0.4 and 0.7. However, on a standards-referenced assessment such as the Montana CRT, it may be appropriate to include some items with very low or very high item difficulty values to ensure sufficient content coverage.

A desirable characteristic of an item is for higher-ability students to perform better on the item than lower-ability students do. The correlation between student performance on a single item and total test score is a commonly used measure of this characteristic of the item. Within classical test theory, the item-test correlation is referred to as the item's discrimination, because it indicates the extent to which successful performance on an item discriminates between high and low scores on the test. For constructed-response items, the item discrimination index used was the Pearson product-moment correlation; for dichotomous items (multiple-choice and short-answer), the corresponding statistic is commonly referred to as a point-biserial correlation. The theoretical range of these statistics is –1.0 to +1.0, with a typical observed range from 0.2 to 0.6.

Discrimination indices can be thought of as measures of how closely an item assesses the same knowledge and skills assessed by other items contributing to the criterion total score. That is, the discrimination index can be thought of as a measure of construct consistency.

A summary of the item difficulty and item discrimination statistics for each subject/grade combination is presented in Table 6-1. Note that the statistics are presented for all items as well as by item type (multiple-choice and open-response, which includes both constructed-response and short-answer items). The mean difficulty and discrimination values shown in the table are within generally acceptable and expected ranges.

Table 6-1. 2009–10 MontCAS: Summary of Item Difficulty and Discrimination Statistics by Subject and Grade

Subject	*Grade*	*Item type*	*Number of items*	p-*value*		*Discrimination*	
				Mean	*Standard deviation*	*Mean*	*Standard deviation*
Mathematics	3	All	60	0.70	0.17	0.37	0.08
		MC	55	0.70	0.17	0.36	0.08
		OR	5	0.67	0.17	0.44	0.06
	4	All	60	0.66	0.15	0.38	0.09
		MC	55	0.67	0.15	0.37	0.08
		OR	5	0.57	0.07	0.51	0.11
	5	All	60	0.64	0.15	0.38	0.09
		MC	55	0.64	0.15	0.38	0.09
		OR	5	0.59	0.20	0.46	0.11
	6	All	60	0.60	0.15	0.38	0.11
		MC	55	0.61	0.14	0.37	0.10
		OR	5	0.48	0.23	0.51	0.13
	7	All	60	0.59	0.16	0.40	0.10
		MC	55	0.59	0.16	0.39	0.08
		OR	5	0.55	0.14	0.58	0.07
	8	All	58	0.57	0.17	0.35	0.09
		MC	54	0.57	0.17	0.33	0.08

continued

Subject	Grade	Item type	Number of items	p-value		Discrimination	
				Mean	Standard deviation	Mean	Standard deviation
Mathematics	8	OR	4	0.50	0.07	0.52	0.05
	10	All	60	0.51	0.16	0.35	0.10
		MC	55	0.52	0.16	0.34	0.09
		OR	5	0.37	0.11	0.47	0.10
Reading	3	All	54	0.70	0.15	0.39	0.07
		MC	52	0.71	0.13	0.38	0.07
		OR	2	0.31	0.03	0.50	0.01
	4	All	54	0.70	0.12	0.39	0.07
		MC	52	0.71	0.10	0.39	0.07
		OR	2	0.36	0.01	0.46	0.01
	5	All	54	0.72	0.12	0.38	0.07
		MC	52	0.73	0.11	0.38	0.07
		OR	2	0.48	0.01	0.46	0.08
	6	All	54	0.71	0.12	0.38	0.07
		MC	52	0.72	0.12	0.38	0.06
		OR	2	0.51	0.03	0.47	0.01
	7	All	54	0.71	0.11	0.41	0.07
		MC	52	0.72	0.10	0.41	0.07
		OR	2	0.50	0.04	0.55	0.04
	8	All	54	0.74	0.08	0.39	0.08
		MC	52	0.74	0.08	0.39	0.08
		OR	2	0.54	0.04	0.48	0.06
	10	All	54	0.73	0.10	0.40	0.07
		MC	52	0.74	0.10	0.39	0.06
		OR	2	0.50	0.06	0.55	0.03
Science	4	All	55	0.66	0.14	0.31	0.08
		MC	53	0.66	0.14	0.30	0.07
		OR	2	0.49	0.23	0.45	0.01
	8	All	55	0.60	0.15	0.34	0.08
		MC	53	0.60	0.15	0.33	0.08
		OR	2	0.62	0.13	0.52	0.02
	10	All	55	0.58	0.14	0.34	0.09
		MC	53	0.60	0.12	0.34	0.08
		OR	2	0.31	0.21	0.54	0.06

MC = multiple-choice; OR = open-response

A comparison of indices across grade levels is complicated because these indices are population dependent. Direct comparisons would require that either the items or students were common across groups. Since that is not the case, it cannot be determined whether differences in performance across grade levels are due to differences in student abilities, differences in item difficulties, or both. With this caveat in mind, it appears generally that, for mathematics and, to a lesser extent, science, students in higher grades found their items more difficult than students in lower grades found theirs, while in reading, the difficulty values are fairly constant across grades.

Comparing the difficulty indices of multiple-choice items and constructed-response or short-answer items is inappropriate because multiple-choice items can be answered correctly by guessing. Thus, it is not surprising that the difficulty indices for multiple-choice items tend to be higher (indicating that students performed better on these items) than the difficulty indices for constructed-response items. Similarly,

discrimination indices for the four-point constructed-response items were larger than those for the dichotomous items due to the greater variability of the former (i.e., the partial credit these items allow) and the tendency for correlation coefficients to be higher given greater variances of the correlates.

In addition to the item difficulty and discrimination summaries presented above, item level classical statistics and item level score distributions were also calculated. Item level classical statistics are provided in Appendix D; item difficulty and discrimination values are presented for each item. The item difficulty and discrimination indices are within generally acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that students who performed well on individual items tended to perform well overall. There were a small number of items with near-zero discrimination indices, but none were negative. While it is not inappropriate to include items with low discrimination values or with very high or very low item difficulty values to ensure that content is appropriately covered, there were very few such cases on the Montana CRT. Item level score-point distributions are provided for constructed-response items in Appendix E; for each item, the percentage of students who received each score point is presented.

6.2 DIFFERENTIAL ITEM FUNCTIONING

Code of Fair Testing Practices in Education (2004) explicitly states that subgroup differences in performance should be examined when sample sizes permit and that actions should be taken to ensure that differences in performance are due to construct-relevant, rather than irrelevant, factors. *Standards for Educational and Psychological Testing* (AERA et al., 1999) includes similar guidelines. As part of the effort to identify such problems, Montana CRT items were evaluated in terms of differential item functioning (DIF) statistics.

For the Montana CRT, the standardization DIF procedure (Dorans and Kulick, 1986) was employed to evaluate subgroup differences. The standardization DIF procedure is designed to identify items for which subgroups of interest perform differently, beyond the impact of differences in overall achievement. The DIF procedure calculates the difference in item performance for two groups of students (at a time) matched for achievement on the total test. Specifically, average item performance is calculated for students at every total score. Then an overall average is calculated, weighting the total score distribution so that it is the same for the two groups.

When differential performance between two groups occurs on an item (i.e., a DIF index in the "low" or "high" categories, explained below), it may or may not be indicative of item bias. Course-taking patterns or differences in school curricula can lead to DIF but for construct-relevant reasons. On the other hand, if subgroup differences in performance could be traced to differential experience (such as geographical living conditions or access to technology), the inclusion of such items should be reconsidered.

Computed DIF indices have a theoretical range from -1.0 to 1.0 for multiple-choice and short-answer items, and the index is adjusted to the same scale for constructed-response items. Dorans and Holland (1993)

suggested that index values between -0.05 and 0.05 should be considered negligible. The preponderance of Montana CRT items fell within this range. Dorans and Holland further stated that items with values between -0.10 and -0.05 and between 0.05 and 0.10 (i.e., "low" DIF) should be inspected to ensure that no possible effect is overlooked, and that items with values outside the [-0.10, 0.10] range (i.e., "high" DIF) are more unusual and should be examined very carefully.[2]

For the 2009–10 Montana CRT, five subgroup comparisons were evaluated for DIF:

- Male versus female
- White versus Native American
- White versus Hispanic
- Disability versus no disability
- Low income versus not low income

Other race/ethnicity groups (e.g., Asians or African Americans) were not analyzed using DIF procedures, because limited sample sizes would have inflated type I error rates. The tables in Appendix F present the number of items classified as either "low" or "high" DIF, overall and by group favored.

6.3 DIMENSIONALITY ANALYSIS

Because tests are constructed with multiple content area subcategories and their associated knowledge and skills, the potential exists for a large number of dimensions being invoked beyond the common primary dimension. Generally, the subcategories are highly correlated with each other; therefore, the primary dimension they share typically explains an overwhelming majority of variance in test scores. In fact, the presence of just such a dominant primary dimension is the psychometric assumption that provides the foundation for the unidimensional IRT models that are used for calibrating, linking, scaling, and equating the 2009–10 MontCAS test forms.

The purpose of dimensionality analyses is to investigate whether violation of the assumption of test unidimensionality is statistically detectable and, if so, (a) the degree to which unidimensionality is violated and (b) the nature of the multidimensionality. Findings from dimensionality analyses performed on the 2009–10 MontCAS common items for mathematics, reading, and science are reported below. (Note: only common items were analyzed since they are used for score reporting.)

The dimensionality analyses were conducted using the nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout, Froelich, & Gao, 2001) and DETECT (Zhang & Stout, 1999). Both of these methods use as their basic statistical building block the estimated average conditional covariances for item pairs. A

[2] It should be pointed out here that DIF for items is evaluated initially at the time of field testing. If an item displays high DIF, it is flagged for review by a Measured Progress content specialist. The content specialist consults with the Department to determine whether to include the flagged item in a future operational test administration.

conditional covariance is the covariance between two items conditioned on total score for the rest of the test, and the average conditional covariance is obtained by averaging over all possible conditioning scores. When a test is strictly unidimensional, all conditional covariances are expected to take on values within random noise of zero, indicating statistically independent item responses for examinees with equal expected scores. Nonzero conditional covariances are essentially violations of the principle of local independence, and local *dependence* implies multidimensionality. Thus, nonrandom patterns of positive and negative conditional covariances are indicative of multidimensionality.

DIMTEST is a hypothesis-testing procedure for detecting violations of local independence. The data are first randomly divided into a training sample and a cross-validation sample. Then an exploratory analysis of the conditional covariances is conducted on the training sample data to find the cluster of items that displays the greatest evidence of local dependence. The cross-validation sample is then used to test whether the conditional covariances of the selected cluster of items displays local dependence, conditioning on total score on the non-clustered items. The DIMTEST statistic follows a standard normal distribution under the null hypothesis of unidimensionality.

DETECT is an effect-size measure of multidimensionality. As with DIMTEST, the data are first randomly divided into a training sample and a cross-validation sample (these samples are drawn independent of those used with DIMTEST). The training sample is used to find a set of mutually exclusive and collectively exhaustive clusters of items that best fit a systematic pattern of positive conditional covariances for pairs of items from the same cluster and negative conditional covariances from different clusters. Next, the clusters from the training sample are used with the cross-validation sample data to average the conditional covariances: within-cluster conditional covariances are summed, from this sum the between-cluster conditional covariances are subtracted, this difference is divided by the total number of item pairs, and this average is multiplied by 100 to yield an index of the average violation of local independence for an item pair. DETECT values less than 0.2 indicate very weak multidimensionality (or near unidimensionality), values of 0.2 to 0.4 weak to moderate multidimensionality, values of 0.4 to 1.0 moderate to strong multidimensionality, and values greater than 1.0 very strong multidimensionality.

DIMTEST and DETECT were applied to the 2009–10 MontCAS. The data for each grade and content area were split into a training sample and a cross-validation sample as described above. Every grade/content area combination included at least 10,000 student examinees, so every training sample and cross-validation sample included at least 5,000 students. DIMTEST was then applied to every grade/content area combination. DETECT was applied to each dataset for which the DIMTEST null hypothesis was rejected in order to estimate the effect size of the multidimensionality.

Because of the large sample sizes for the Montana tests, DIMTEST would be sensitive even to quite small violations of unidimensionality, and the null hypothesis was strongly rejected for every dataset with all p-values being less than 0.00005. Strong rejection of the null hypothesis of unidimensionality is not surprising because strict unidimensionality is an idealization that almost never holds exactly for a given

dataset. Thus, it was important to use DETECT to estimate the effect size of the violations of local independence found by DIMTEST. Table 6-2 displays the multidimensional effect size estimates from DETECT.

Table 6-2. 2009–10 MontCAS: Multidimensionality Effect Sizes by Subject and Grade

Subject	*Grade*	*Multidimensionality effect size*	
		2008–09	*2009–10*
Mathematics	3	0.11	0.08
	4	0.12	0.13
	5	0.13	0.13
	6	0.12	0.15
	7	0.13	0.14
	8	0.12	0.11
	10	0.11	0.12
Reading	3	0.12	0.10
	4	0.12	0.11
	5	0.10	0.09
	6	0.12	0.09
	7	0.08	0.10
	8	0.15	0.10
	10	0.10	0.14
Science	4	0.11	0.11
	8	0.09	0.18
	10	0.09	0.12

All the DETECT values for 2009–10 indicated very weak multidimensionality. The average DETECT values for the three content areas were 0.12 for Math, 0.10 for Reading, and 0.14 for Science. Also shown in Table 6-2 are the values reported in last year's dimensionality analyses. The DETECT indices for the individual content areas for each grade are seen to be very similar between the two years. In particular, both sets of values indicate very weak multidimensionality for all the tests, and, consequently, the averages for the three content areas for 2009–10 (0.12 for Math, 0.11 for Reading, and 0.10 for Science) are similar to the 2008–09 averages. We also investigated how DETECT divided the tests into clusters to see if there were any discernable patterns with respect to item type (i.e., multiple-choice and constructed-response) or other factors. Only Science in grade 4 had a DETECT cluster that substantially separated the multiple-choice (MC) items from the constructed-response (CR) items. This lack of separation of multiple-choice and constructed-response items also occurred in the 2008–09 and in the 2007–08 tests. A more thorough investigation employing experts in the substantive content of the test forms may result in identification of clusters related to the skills and knowledge areas measured by the items. In any case the violations of local independence from all such effects, as evidenced by the DETECT effect sizes, were very small and do not warrant any changes in test design or scoring.

CHAPTER 7. IRT SCALING AND EQUATING

In addition to the classical test theory item analyses previously described, the Montana CRT was analyzed according to IRT models. IRT analyses were first used to place all 2009–10 forms on the same scale, and then to equate the 2009–10 test to the previous year's test. Details on the IRT calibration and equating procedures for the Montana CRT are described below.

7.1 ITEM RESPONSE THEORY

IRT uses mathematical models to define a relationship between an unobserved measure of student performance, usually referred to as theta (θ), and the probability (p) of getting a dichotomous item correct or of getting a particular score on a polytomous item. In IRT, it is assumed that all items are independent measures of the same construct (i.e., of the same θ). Another way to think of θ is as a mathematical representation of the latent trait of interest. Several common IRT models are used to specify the relationship between θ and p (Hambleton and van der Linden, 1997; Hambleton and Swaminathan, 1985). The process of determining the specific mathematical relationship between θ and p is called item calibration. After items are calibrated, they are defined by a set of parameters that specify a nonlinear, monotonically increasing relationship between θ and p. For the Montana CRT, the one parameter logistic (1PL) model was used for multiple-choice and short-answer items and the partial credit model was used for the constructed-response items.

For polytomous items, the generalized partial credit model can be defined as:

$$P_{jk}(\theta) = \frac{\exp \sum_{v=0}^{k} \left[Da_j \left(\theta - b_j + d_v \right) \right]}{\sum_{c=1}^{m} \exp \sum_{v=1}^{c} \left[Da_j \left(\theta - b_j + d_v \right) \right]}$$

where
j indexes items,
k indexes students,
a represents item discrimination,
b represents item difficulty,
d represents category step parameter, and
D is a normalizing constant equal to 1.701.

In the case of the Montana CRT, the a_j term in the equation is equal to 1.0 for all items. The equation reduces to the following for dichotomous items with no step parameters (d_v):

$$P_j(\theta) = \frac{\exp D\left(\theta - b_j\right)}{1 + \exp D\left(\theta - b_j\right)}$$

For more information about item calibration and determination, the reader is referred to Lord and Novick (1968), Hambleton and Swaminathan (1985), or Baker and Kim (2004).

7.2 ITEM RESPONSE RESULTS

PARSCALE v3.5 (Muraki & Bock, 1999) software was used to perform all IRT analyses for the Montana CRT. Each item occupied only one block in the calibration run, and the 1.701 normalizing constant was used. A default convergence criterion of 0.001 was used. The tables in Appendix G give the IRT item parameters of all dichotomous (multiple-choice and short-answer) and polytomous (constructed-response) items on the 2009–10 Montana CRT tests by subject and grade.

Appendix H provides the test characteristic curves (TCCs) and test information functions (TIFs). TCCs display the expected (average) raw score associated with each θ_j value between -4.0 and 4.0. Mathematically, the TCC is computed by summing the item characteristic curves (ICCs) of all items that contribute to the raw score. The expected raw score at a given value of θ_j is

$$E(X \mid \theta_j) = \sum_{i=1}^{n} P_i\left(1 \middle| \theta_j\right),$$

where
i indexes the items (and n is the number of items contributing to the raw score),
j indexes students (here, θ_j runs from –4 to 4), and
$E(X \mid \theta_j)$ is the expected raw score for a student of ability θ_j.

The expected raw score monotonically increases with θ_j, consistent with the notion that students of high ability tend to earn higher raw scores than do students of low ability. Most TCCs are "S-shaped," flatter at the ends of the distribution and steeper in the middle.

The TIF displays the amount of statistical information that the test provides at each value of θ_j. Information functions depict test precision across the entire latent trait continuum. There is an inverse relationship between the information of a test and its standard error of measurement (SEM). For long tests, the SEM at a given θ_j is approximately equal to the inverse of the square root of the statistical information at θ_j (Hambleton, Swaminathan, & Rogers, 1991), as follows:

$$SEM(\theta_j) = \frac{1}{\sqrt{I(\theta_j)}}$$

Compared to the tails, TIFs are often higher near the middle of the θ distribution where most students are located and where most items are sensitive by design.

7.3 EQUATING

The purpose of equating is to ensure that scores obtained from different forms of a test are equivalent to each other. Equating may be used if multiple test forms are administered in the same year, as well as to equate one year's forms to those given in the previous year. Equating ensures that students are not given an unfair advantage or disadvantage because the test form they took is easier or harder than those taken by other students.

Equating for the Montana CRT used the *anchor-test-nonequivalent-groups* design described by Petersen, Kolen, & Hoover (1989). In this equating design, no assumption is made about the equivalence of the examinee groups taking different test forms (that is, naturally occurring groups are assumed). IRT is particularly useful for equating nonequivalent groups (Allen & Yen, 1979).The fixed common-item IRT procedure was used: The anchor items from the previous year's administration were identified during this year's calibrations, and their IRT parameters were fixed to last year's values. This method results in all person and item parameters being on the same θ scale as they were in the previous year. The procedures used for equating and scaling do not change the rank ordering of students, give more weight to particular items, or change students' performance-level classifications.

7.4 EQUATING RESULTS

An Equating Report was submitted to the OPI for their approval prior to production of student reports. The equating report details the results of a variety of quality control activities that were implemented within the Psychometrics and Research Department during IRT calibration and equating, including examining *b* plots and TCCs and conducting delta and rescore analyses. The evaluations of the equating results are summarized below.

The number of Newton cycles required for convergence for each subject and grade during the IRT analysis can be found in Table 7-1. The number of cycles required in order for the solution to converge fell within acceptable ranges.

Table 7-1. 2009–10 Montana CRT: Number of Newton Cycles Required for Convergence

Subject	*Grade*	*Cycles*
Mathematics	3	34
	4	25
	5	19
	6	10
	7	11
	8	4
	10	5
Reading	3	33
	4	34
	5	36
	6	33

continued

Subject	Grade	Cycles
Reading	7	38
	8	39
	10	40
Science	4	5
	8	5
	10	5

Appendix I presents the results from the delta analysis. This procedure was used to evaluate the performance of equating items, and the discard status presented in the appendix indicates whether the item was used in equating. As can be seen in the appendix, as well as in Table 7-2 below, a very small number of items were identified as problematic based on the results of the delta analyses and were excluded from use in equating.

Also presented in Appendix I are the results from the rescore analysis. With this analysis, 200 random papers from the previous year were interspersed with this year's papers to evaluate scorer consistency from one year to the next. All effect sizes were well below the criterion value for excluding an item as an equating item, 0.80 (in absolute value).

Table 7-2 below shows all items that required intervention during IRT calibration and equating. As can be seen in the table, all items on the watch list were identified as a result of the delta analyses. In all cases, the identified item was excluded from use in equating.

Table 7-2. 2009–10 MontCAS: Items that Required Intervention During IRT Calibration and Equating

IREF	Subject	Grade	Reasons	Action
242889	MAT	4	Delta Analysis	Removed from equating
77340	MAT	6	Delta Analysis	Removed from equating
88363	MAT	8	Delta Analysis	Removed from equating
43743	MAT	10	Delta Analysis	Removed from equating
92791	REA	3	Delta Analysis	Removed from equating
92766	REA	3	Delta Analysis	Removed from equating
93444	REA	5	Delta Analysis	Removed from equating
95863	REA	8	Delta Analysis	Removed from equating
95187	REA	10	Delta Analysis	Removed from equating
89277	SCI	8	Delta Analysis	Removed from equating
75706	SCI	10	Delta Analysis	Removed from equating

7.5 ACHIEVEMENT STANDARDS

Cutpoints for the Montana CRT in reading and mathematics were set at standard-setting meetings held in June and July 2006 and for the Montana CRT in science in June 2008. Details of the standard-setting procedures can be found in the standard-setting reports and technical reports of those years. The cuts on the theta scale that were established at those meetings are presented in Table 7-3 below. The θ-metric cut scores that emerged from the standard setting meetings will remain fixed throughout the assessment program unless

standards are reset for any reason. Also shown in the table are the cutpoints on the reporting score scale (described below).

Table 7-3. 2009–10 MontCAS: Cut Scores on the Theta Metric and Reporting Scale by Subject and Grade

Subject	Grade	Theta			Scaled score				
		Cut 1	Cut 2	Cut 3	Minimum	Cut 1	Cut 2	Cut 3	Maximum
Mathematics	3	-0.54340	-0.20337	0.44500	200	225	250	290	300
	4	-0.29081	0.05530	0.65734	200	225	250	291	300
	5	-0.55315	-0.20313	0.38248	200	225	250	289	300
	6	-0.55054	-0.17902	0.36958	200	225	250	287	300
	7	-0.51684	-0.16514	0.35144	200	225	250	289	300
	8	-0.52251	-0.09914	0.46022	200	225	250	283	300
	10	-0.57541	-0.06623	0.50451	200	225	250	281	300
Reading	3	-1.03019	-0.52098	0.26228	200	225	250	287	300
	4	-0.64979	-0.19215	0.55362	200	225	250	289	300
	5	-0.86117	-0.43483	0.24763	200	225	250	287	300
	6	-0.82220	-0.42340	0.26115	200	225	250	289	300
	7	-0.87767	-0.44082	0.29929	200	225	250	288	300
	8	-0.54622	-0.17634	0.50092	200	225	250	289	300
	10	-0.42862	-0.08340	0.55241	200	225	250	289	300
Science	4	-0.70081	-0.14474	0.55956	200	225	250	282	300
	8	-0.57275	-0.07715	0.58285	200	225	250	283	300
	10	-0.37793	0.12744	0.52244	200	225	250	270	300

7.5.2 Distributions

Table J-1 in Appendix J shows performance level distributions for each of the last three years by subject and grade.

7.6 SCALED SCORES

7.6.1 Description of Scale

Montana CRT scores in each content area are reported on a scale ranging from 200 to 300. By providing information that is more specific about the position of a student's results, scaled scores supplement performance level scores. School- and district-level scaled scores are calculated by computing the average of student-level scaled scores. Students' raw scores (i.e., total number of points) on the 2009–10 Montana CRT were translated to scaled scores using a data analysis process called *scaling*. Scaling simply converts from one scale to another. In the same way that a given temperature can be expressed on either Fahrenheit or Celsius scales, or the same distance can be expressed in either miles or kilometers, student scores on the 2009–10 Montana CRT tests can be expressed in raw or scaled scores.

It is important to note that converting from raw scores to scaled scores does not change students' performance-level classifications. Given the relative simplicity of raw scores, it is fair to ask why scaled scores instead of raw scores are used in Montana CRT reports. Foremost, scaled scores offer the advantage of

simplifying result reporting across content areas, grade levels, and subsequent years. Because the standard-setting process typically results in different cut scores across content areas on a raw score basis, it is useful to transform these raw cut scores to a scale that is more easily interpretable and consistent. For the Montana CRT, a score of 225 is the cut score between the Novice and Nearing Proficiency performance levels. This is true regardless of content area, grade level, or year. For example, the raw cut score between Novice and Nearing Proficiency may be 35 in grade 8 mathematics but may be 33 in grade 10 mathematics. Using scaled scores greatly simplifies the task of understanding how a student performed. The raw score to scaled score look-up tables for each content area and grade are presented in Appendix K.

7.6.2 Calculations

For Montana CRT, scaled scores were obtained by a simple translation of students' scores using a linear equation of the form

$$SS = mY + b$$

where
m is the slope,
b is the intercept, and
Y represents the student's score.

A separate linear transformation was used for each grade/content area combination. Each line was determined by using threshold values obtained via standard setting and fixing the Novice/Nearing Proficiency and Nearing Proficiency/Proficient scaled score cuts to 225 and 250, respectively. The cut between Proficient and Advanced was then allowed to vary across grades and content areas. The scaled score values obtained using this formula were rounded to the nearest integer and truncated, as necessary, so that no student received a score lower than 200 or higher than 300.

For science, the student score used for scaling was the ability estimate on the theta scale, $\hat{\theta}$, which was found from the students' raw scores by mapping through the TCC. For reading and mathematics, on the other hand, scaling was done from raw score. As with science, the students' raw scores on the 2009–10 test were transformed into ability estimates on the theta scale using the TCC. These ability estimates were then transformed into an expected raw score on the reference test form (2005–06, when standards were established for reading and mathematics) using the TCC for the reference test. This expected raw score was then scaled onto the reporting metric.

Table 7-4 shows the scaling constants by subject and grade.

Table 7-4. 2009–10 MontCAS: Scaled Score Slope and Intercept by Subject and Grade

Subject	*Grade*	*Slope*	*Intercept*
Mathematics	3	3.1692	118.5242
	4	3.0431	141.4551
	5	2.8083	155.7965

continued

Subject	*Grade*	*Slope*	*Intercept*
Mathematics	6	2.7906	159.5450
	7	3.0378	159.7850
	8	2.4365	172.4985
	10	2.0947	181.1735
Reading	3	2.4370	182.0623
	4	2.5939	174.3429
	5	2.7798	161.4892
	6	3.0026	154.7492
	7	2.5872	169.9388
	8	3.0898	145.1710
	10	3.1680	130.2323
Science	4	44.9584	256.5073
	8	50.4439	253.8917
	10	49.4687	243.6957

7.6.3 Distributions

Graphs of the scaled score cumulative frequency distributions for the last three years are presented in Appendix J. Note that the graphs show the percent of students at or below each scaled score, thus the lowest line in a given graph depicts the highest performing group. For example, in the graph for grade 3 mathematics (Figure L-1), the line showing the cumulative distribution for 2009–10 is consistently lower than that for 2008–09 which, in turn, is consistently lower than that for 2007–08. This pattern indicates that student performance on the grade 3 mathematics test has improved in each of the last two years.

CHAPTER 8. RELIABILITY

Although an individual item's performance is an important focus for evaluation, a complete evaluation of an assessment must also address the way items function together and complement one another. Tests that function well provide a dependable assessment of the student's level of ability. Unfortunately, no test can do this perfectly. A variety of factors can contribute to a given student's score being either higher or lower than his or her true ability. For example, a student may misread an item, or mistakenly fill in the wrong bubble when he or she knew the answer. Collectively, extraneous factors that impact a student's score are referred to as measurement error. Any assessment includes some amount of measurement error; that is, no measurement is perfect. This is true of all academic assessments—some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. When tests have a high amount of measurement error, student scores are very unstable. Students with high ability may get low scores or vice versa. Consequently, one cannot reliably measure a student's true level of ability with such a test. Assessments that have less measurement error (i.e., errors made are small on average and student scores on such a test will consistently represent their ability) are described as reliable.

There are a number of ways to estimate an assessment's reliability. One possible approach is to give the same test to the same students at two different points in time. If students receive the same scores on each test, then the extraneous factors affecting performance are small and the test is reliable. (This is referred to as "test-retest reliability.") A potential problem with this approach is that students may remember items from the first administration or may have gained (or lost) knowledge or skills in the interim between the two administrations. A solution to the "remembering items" problem is to give a different, but parallel test at the second administration. If student scores on each test correlate highly the test is considered reliable. (This is known as "alternate forms reliability," because an alternate form of the test is used in each administration.) This approach, however, does not address the problem that students may have gained (or lost) knowledge or skills in the interim between the two administrations. In addition, the practical challenges of developing and administering parallel forms generally preclude the use of parallel forms reliability indices. One way to address the latter problems is to split the test in half and then correlate students' scores on the two half-tests; this in effect treats each half-test as a complete test. By doing this, the problems associated with an intervening time interval and of creating and administering two parallel forms of the test are alleviated. This is known as a "split-half estimate of reliability." If the two half-test scores correlate highly, items on the two half-tests must be measuring very similar knowledge or skills. This is evidence that the items complement one another and function well as a group. This also suggests that measurement error will be minimal.

The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation, since each different possible split of the test halves will result in a different correlation. Another problem with the split-half method of calculating reliability is that it underestimates reliability, because test length is cut in half. All else being equal, a shorter

test is less reliable than a longer test. Cronbach (1951) provided a statistic, α (alpha), which eliminates the problem of the split-half method by comparing individual item variances to total test variance. Cronbach's α was used to assess the reliability of the 2009–10 Montana CRT:

$$\alpha \equiv \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma^2_{(Y_i)}}{\sigma^2_x}\right]$$

where
i indexes the item,
n is the total number of items,
$\sigma^2_{(Y_i)}$ represents individual item variance, and
σ^2_x represents the total test variance.

8.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT

Table 8-1 presents descriptive statistics, Cronbach's α coefficient, and raw score standard errors of measurement (SEMs) for each content area and grade. (Statistics are based on common items only.)

Table 8-1. 2009–10 MontCAS: Raw Score Descriptive Statistics, Cronbach's Alpha and Standard Errors of Measurement (SEM) by Subject and Grade

Subject	*Grade*	*Number of students*	*Raw score*			*Alpha*	*SEM*
			Maximum	*Mean*	*Standard deviation*		
Mathematics	3	10544	66	45.13	10.94	0.91	3.28
	4	10262	66	42.76	12.47	0.91	3.69
	5	10514	66	41.22	12.47	0.91	3.65
	6	10442	66	37.73	12.67	0.92	3.67
	7	10457	66	37.62	13.41	0.92	3.70
	8	10629	61	34.09	10.98	0.90	3.53
	10	10389	66	32.10	11.95	0.90	3.69
Reading	3	10509	60	39.39	10.29	0.91	3.11
	4	10225	60	40.01	10.63	0.91	3.13
	5	10491	60	41.68	10.15	0.91	3.09
	6	10419	60	41.37	10.32	0.91	3.17
	7	10460	60	41.58	10.93	0.92	3.10
	8	10637	60	43.05	10.56	0.91	3.13
	10	10422	60	42.63	10.48	0.91	3.07
Science	4	10265	61	38.94	9.42	0.86	3.48
	8	10671	61	36.96	10.75	0.88	3.70
	10	10450	61	33.99	10.65	0.89	3.52

For mathematics, the reliability coefficients ranged from 0.90 to 0.92; for reading, from 0.91 to 0.92; and, for science, from 0.86 to 0.89. Because different grades and content areas have different test designs

(e.g., the number of items varies by test), it is inappropriate to make inferences about the quality of one test by comparing its reliability to that of another test from a different grade and/or content area.

8.2 2009–10 Subgroup Reliability

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2009–10 Montana CRT. Appendix L presents reliabilities for various subgroups of interest. Subgroup Cronbach's α's were calculated using the formula defined above based only on the members of the subgroup in question in the computations; values are only calculated for subgroups with 10 or more students. For mathematics, subgroup reliabilities ranged from 0.67 to 0.94; for reading, from 0.79 to 0.94; and for science from 0.73 to 0.91.

For several reasons, the results of this section should be interpreted with caution. First, inherent differences between grades and content areas preclude making valid inferences about the quality of a test based on statistical comparisons with other tests. Second, reliabilities are dependent not only on the measurement properties of a test but on the statistical distribution of the studied subgroup. For example, it can be readily seen in Appendix L that subgroup sample sizes may vary considerably, which results in natural variation in reliability coefficients. Or α, which is a type of correlation coefficient, may be artificially depressed for subgroups with little variability (Draper & Smith, 1998). Third, there is no industry standard to interpret the strength of a reliability coefficient, and this is particularly true when the population of interest is a single subgroup.

8.3 Reporting Subcategory Reliability

Of even more interest are reliabilities for the reporting subcategories within Montana CRT content areas, described in Chapter 3. Cronbach's α coefficients for subcategories were calculated via the same formula defined previously using just the items of a given subcategory in the computations. Results are presented in Appendix L. Once again as expected, because they are based on a subset of items rather than the full test, computed subcategory reliabilities were lower (sometimes substantially so) than were overall test reliabilities, and interpretations should take this into account.

For mathematics, subcategory reliabilities ranged from 0.39 to 0.82; for reading, from 0.54 to 0.81; and for science, from 0.27 to 0.68. The subcategory reliabilities were lower than those based on the total test and approximately to the degree one would expect based on classical test theory. Qualitative differences between grades and content areas once again preclude valid inferences about the quality of the full test based on statistical comparisons among subtests.

8.4 Interrater Consistency

Chapter 5 of this report describes in detail the processes that were implemented to monitor the quality of the hand-scoring of student responses for short-answer and constructed-response items. One of these

processes was double-blind scoring: approximately 2% of student responses were randomly selected and scored independently by two different scorers. Results of the double-blind scoring were used during scoring to identify scorers who required retraining or other intervention and are presented here as evidence of the reliability of the Montana CRT. A summary of the interrater consistency results are presented in Table 8-2 below. Results in the table are collapsed across the hand-scored items by subject, grade, and number of score categories (two for short-answer items and five for constructed-response items). The table shows the number of included scores, the percent exact agreement, percent adjacent agreement, correlation between the first two sets of scores, and the percent of responses that required a third score. This same information is provided at the item level in Appendix M.

Table 8-2. Summary of Interrater Consistency Statistics Collapsed Across Items by Subject and Grade

Subject	*Grade*	*Number of score categories*	*Number of included scores*	*Percent exact*	*Percent adjacent*	*Correlation*	*Percent of third scores*
Mathematics	3	2	639	99.37	0.63	0.98	0.00
		5	422	91.94	7.58	0.96	0.47
	4	2	638	99.06	0.94	0.98	0.00
		5	432	88.66	9.49	0.96	1.85
	5	2	636	98.74	1.26	0.97	0.00
		5	418	85.89	12.20	0.95	1.91
	6	2	641	97.50	2.50	0.95	0.00
		5	423	79.67	17.73	0.92	2.60
	7	2	631	98.26	1.74	0.96	0.00
		5	410	83.17	15.12	0.94	1.46
	8	2	640	98.75	1.25	0.97	0.00
		5	207	75.36	21.74	0.91	2.90
	10	2	617	95.95	4.05	0.92	0.00
		5	401	71.32	24.69	0.86	2.99
Reading	3	5	428	66.36	31.31	0.73	2.10
	4	5	422	65.17	33.89	0.78	0.95
	5	5	428	67.76	31.07	0.77	1.17
	6	5	430	72.56	25.81	0.83	1.40
	7	5	417	61.63	37.17	0.76	1.20
	8	5	426	56.34	40.38	0.75	2.35
	10	5	417	64.99	34.05	0.77	0.96
Science	4	5	401	85.04	13.97	0.94	1.00
	8	5	440	72.05	22.95	0.89	5.00
	10	5	428	72.43	24.53	0.88	3.27

8.5 RELIABILITY OF PERFORMANCE LEVEL CATEGORIZATION

While related to reliability, the accuracy and consistency of classifying students into performance categories are even more important statistics in a standards based reporting framework (Livingston & Lewis, 1995). After the performance levels were specified and students were classified into those levels, empirical analyses were conducted to determine the statistical accuracy and consistency of the classifications. For the Montana CRT, students are classified into one of four performance levels: Novice (N), Nearing Proficiency

(NP), Proficient (P), or Advanced (A). This section of the report explains the methodologies used to assess the reliability of classification decisions, and results are given.

Accuracy refers to the extent to which decisions based on test scores match decisions that would have been made if the scores did not contain any measurement error. Accuracy must be estimated, because errorless test scores do not exist. Consistency measures the extent to which classification decisions based on test scores match the decisions based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are given to the same group of students. In operational test programs, however, such a design is usually impractical. Instead, techniques have been developed to estimate both the accuracy and consistency of classification decisions based on a single administration of a test. The Livingston and Lewis (1995) technique was used for the 2009–10 Montana CRT because it is easily adaptable to all types of testing formats, including mixed format tests.

The accuracy and consistency estimates reported in Appendix N make use of "true scores" in the classical test theory sense. A true score is the score that would be obtained if a test had no measurement error. Of course, true scores cannot be observed and so must be estimated. In the Livingston and Lewis method, estimated true scores are used to categorize students into their "true" classifications.

For the 2009–10 Montana CRT, after various technical adjustments (described in Livingston & Lewis, 1995), a four by four contingency table of accuracy was created for each content area and grade, where cell $[i, j]$ represented the estimated proportion of students whose true score fell into classification i (where i = 1 to 4) and observed score into classification j (where j = 1 to 4). The sum of the diagonal entries (i.e., the proportion of students whose true and observed classifications matched) signified overall accuracy.

To calculate consistency, true scores were used to estimate the joint distribution of classifications on two independent, parallel test forms. Following statistical adjustments per Livingston and Lewis (1995), a new four by four contingency table was created for each content area and grade and populated by the proportion of students who would be categorized into each combination of classifications according to the two (hypothetical) parallel test forms. Cell $[i, j]$ of this table represented the estimated proportion of students whose observed score on the first form would fall into classification i (where i = 1 to 4) and whose observed score on the second form would fall into classification j (where j = 1 to 4). The sum of the diagonal entries (i.e., the proportion of students categorized by the two forms into exactly the same classification) signified overall consistency.

Another way to measure consistency is to use Cohen's (1960) coefficient κ (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. It is calculated using the following formula:

$$\kappa = \frac{\text{(Observed agreement)} \ - \ \text{(Chance agreement)}}{1 \ - \ \text{(Chance agreement)}} = \frac{\sum_i C_{ii} - \sum_i C_{i.}C_{.i}}{1 - \sum_i C_{i.}C_{.i}},$$

where
$C_{i.}$ is the proportion of students whose observed performance level would be Level i (where i = 1–4) on the first hypothetical parallel form of the test;
C_{i} is the proportion of students whose observed performance level would be Level i (where i = 1–4) on the second hypothetical parallel form of the test;
C_{ii} is the proportion of students whose observed performance level would be Level i (where i = 1–4) on both hypothetical parallel forms of the test.

Because κ is corrected for chance, its values are lower than are other consistency estimates.

8.5.1 Accuracy and Consistency Results

The accuracy and consistency analyses described above are provided in Table N-1 of Appendix N. The table includes overall accuracy and consistency indices, including kappa. Accuracy and consistency values conditional upon performance level are also given. For these calculations, the denominator is the proportion of students associated with a given performance level. For example, the conditional accuracy value is 0.82 for Novice for mathematics grade 3. This figure indicates that among the students whose true scores placed them in this classification, 82 percent would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.74 indicates that 74 percent of students with observed scores in the Novice level would be expected to score in this classification again if a second, parallel test form were used.

For some testing situations, the greatest concern may be decisions around level thresholds. For example, if a college gave credit to students who achieved an Advanced Placement test score of 4 or 5, but not to students with scores of 1, 2, or 3, one might be interested in the accuracy of the dichotomous decision below-4 versus 4-or-above. For the 2009–10 Montana CRT, Table N-2 in Appendix N provides accuracy and consistency estimates at each cutpoint as well as false positive and false negative decision rates. (A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.)

The above indices are derived from Livingston and Lewis's (1995) method of estimating the accuracy and consistency of classifications. It should be noted that Livingston and Lewis discuss two versions of the accuracy and consistency tables. A standard version performs calculations for forms parallel to the form taken. An "adjusted" version adjusts the results of one form to match the observed score distribution obtained in the data. The tables use the standard version for two reasons: (1) this "unadjusted" version can be considered a smoothing of the data, thereby decreasing the variability of the results; and (2) for results dealing with the consistency of two parallel forms, the unadjusted tables are symmetrical, indicating that the two parallel forms have the same statistical properties. This second reason is consistent with the notion of forms

that are parallel; that is, it is more intuitive and interpretable for two parallel forms to have the same statistical distribution.

Descriptive statistics relating to the decision accuracy and consistency (DAC) of the 2009–10 Montana CRT tests can be derived from Table N-1. For mathematics, overall accuracy ranged from 0.77 to 0.80; overall consistency ranged from 0.68 to 0.72; the kappa statistic ranged from 0.55 to 0.60. For reading, overall accuracy ranged from 0.82 to 0.85; overall consistency ranged from 0.75 to 0.79; the kappa statistic ranged from 0.61 to 0.66. Finally, for science, overall accuracy ranged from 0.74 to 0.78; overall consistency ranged from 0.65 to 0.69; the kappa statistic ranged from 0.53 to 0.54. Note that, as with other methods of evaluating reliability, DAC statistics calculated based on small groups can be expected to be lower than those calculated based on larger groups. For this reason, the values presented in Appendix N should be interpreted with caution. In addition, it is important to remember that it is inappropriate to compare DAC statistics between grades and content areas.

CHAPTER 9. SCORE REPORTING

The Montana CRT is designed to measure student performance against Montana's content standards. Consistent with this purpose, results on the CRT were reported in terms of performance levels that describe student performance in relation to these established state standards. There are four performance levels: Novice, Nearing Proficiency, Proficient, and Advanced. (Performance level distributions are given in Appendix J.) Students receive a separate performance level classification (based on total scaled score) in each content area.

State results were provided to the OPI via a secure Web site. Reading, mathematics, and science reporting data for the 2009–10 Montana CRT were made available to systems and schools online via the Montana Analysis and Reporting System (MARS) on May 28, 2010. Student reports were delivered to schools on September 20, 2010. New in 2010, Student Reports were also posted online to be accessible to schools. System test coordinators and teachers were also provided with copies of the *Guide to the 2010 Criterion-Referenced Test and CRT-Alternate Assessment Reports* to assist them in understanding the connection between the assessment and the classroom. The guide provides information about the assessment and the use of assessment results.

School- and system-level results are reported as the number and percentages of students attaining each performance level at each grade level tested. As described below, decision rules were formulated in early 2010 by the OPI and Measured Progress to identify students who, during the reporting process, were to be excluded from school- and system-level reports. (A copy of these decision rules is included in this report as Appendix A.) New in 2010, state-level summary reports were also produced.

The reports described in the sections that follow are separated into two categories. The first set of reports described is static reports, which are provided online as PDF documents; student reports are also provided on paper. The static reports are:

- Student Report (paper and online)
- School, System, and State Summary Reports (online)

The remaining reports are interactive reports, provided via MARS (see sections 9.3 and 9.4 below):

- Class Roster and Item-Level Reports
- Performance-Level Summary
- Released Items Summary Data
- Longitudinal Data Report

Sample Report Shells are included as Appendix O.

9.1 DECISION RULES

As mentioned above, to ensure that reported results for the 2009–10 Montana CRT are accurate relative to collected data and other pertinent information, a document that delineates analysis and reporting rules was created. These decision rules were observed in the analyses of Montana CRT test data and in reporting the test results. Moreover, these rules are the main reference for quality assurance checks.

The decision rules document used for reporting results of the 2010 administration of the Montana CRT is found in Appendix A.

The rules primarily describe the inclusion/exclusion of students at the school-, system-, and state-levels of aggregation. The document also describes rules as they pertain to individual reports. Finally, it describes the classification of students based on their school type or other information provided by the state through the student demographic file (AIM) or collected on the answer booklet.

9.2 STATIC REPORTS

9.2.1 Student Report

The student report was redesigned this year. The report is now printed on 11in. x 17in. paper with an off-center fold.

The student report is produced for each parent of a student who took or was eligible to take the MontCAS CRT. The report is shipped to systems and posted online for school/system access.

The student report gives the results for each subject tested. At grades 3, 5, 6, and 7, these are reading and mathematics. At grades 4, 8, and 10, these are reading, mathematics, and science. The student reports give the earned performance level and scaled score for each subject. The report also provides a comparison of the student's performance to that of the state as a whole for each subject. The report contains the results for each subject at the content standard level. The number of points earned by the student in each content standard is reported as well as the range of points earned by students who achieve proficiency.

9.2.2 Summary Reports

The summary report is produced at the school, system, and state levels. The report is produced for each content area in the grade level. For grades 3, 5, 6, and 7, the content areas are reading and mathematics. For grades 4, 8, and 10, the content areas are reading, mathematics, and science. The report consists of three sections: Distribution of Scores, Subtest Results, and Results for Subgroups of Students.

The Distribution of Scores section of the report contains a breakdown of the performance of included students (as described in the decision rules document) into different scaled score intervals. The number and percent of students that fall into each scaled score interval is shown. There is an overall percentage reported for students that fall into each of the four performance levels (Novice, Nearing Proficiency, Proficient, and Advanced). In the *School Summary Report,* the calculations are done at the school, system, and state levels.

The *System Summary Report* contains results at the system and state levels. The *State Summary Report* contains only state level results.

The Subtest Results section of the report summarizes the average points earned in the different content standards by included students (as described in the decision rules document) in the school, system, and state. The average points earned are compared to the total possible points for each content standard.

The Results for Subgroups of Students section of the report summarizes the performance of included students (as described in the decision rules document) broken down by various reporting categories. For each reporting category, the number of tested (included) students is reported as well as the percentage of students in each of the four performance levels. In the *School Summary Report* this is reported at the school, system, and state levels. In the *System Summary Report* the data are reported at the system and state levels. In the *State Summary Report* the data are reported at state level only.

The list of reporting categories is as follows:

- All Students
- Gender
- Ethnicity (American Indian or Alaska native, Asian, Hispanic, Black or African American, Native Hawaiian or Other Pacific Islander, White)
- Special Education
- Students with a 504 Plan
- Title I (optional)
- Tested with Standard Accommodation
- Tested with Nonstandard Accommodation
- Alternate Assessment (results are not given for this category on the Montana CRT Summary reports)
- Migrant
- Gifted/Talented
- LEP/ELL
- Former LEP Student
- LEP Student Enrolled for First Time in a U.S. School
- Free/Reduced Lunch

Data are suppressed if there are less than ten tested (included) in a reporting category at a given aggregation level.

The data for the reporting categories were provided by information coded on the students' answer booklets by teachers and/or data supplied by the state through an AIM export. Due to relatively low numbers of students in certain reporting categories, school personnel are advised, under FERPA guidelines, to treat these pages confidentially.

9.3 MONTANA ANALYSIS AND REPORTING SYSTEM (MARS)

Using advanced Web technology, *MARS* gives Montana educators and administrators the ability to filter data based on test year, grade level, content area, standard, and student subgroup. This allows administrators to isolate cross-sections of the results and identify areas of strong or poor performance.

The confidential nature of the data in *MARS* necessitates the strict enforcement of site security. All transmissions are done over Secure Socket Layers (SSL). A system of user role definitions and permissions dictates the scope of access granted to individual users. Organizations (system or school levels) are given administrative power to grant or deny access to their data within the system, and they have the ability to disable users. Personnel using *MARS* may be granted permission to view students' results at an organizational level, or only a select group as defined by the administrator. Predefined reports are included in the system, as is the ability to render and print additional copies.

9.3.1 User Accounts

In *MARS*, principals have the ability to create unique user accounts by assigning specific usernames and passwords to educators in their school such as teachers, curriculum coordinators, or special education coordinators. Once the accounts have been created, individual students may be assigned to each user account. After users have received their usernames and passwords, they are able to log in to their accounts and access the interactive reports which will be populated only with the subgroup of students assigned to them.

Information about the interactive reports and setting up user accounts is available in the *Analysis & Reporting System User Manual* that is available for download on the *MARS* system.

9.4 INTERACTIVE REPORTS

As mentioned above, there are four interactive reports that were available from *MARS*: Item Analysis Report, Performance Level Summary, Released Items Summary Data, and Longitudinal Data. Each of these interactive reports is described in the following sections. Sample interactive reports are provided in Appendix P. To access these four interactive reports, the user clicked the interactive tab on the home page of the system and selected the report desired from the drop down menu. Next, the user applied basic filtering options, such as the name of the district or school and the grade level/content area test, to open the specific report. At this point, the user had the option of printing the report for the entire grade level or applying advanced filtering options to select a subgroup of students to analyze. Advanced filtering options include gender, ethnicity,

limited English proficient (LEP), IEP, low income, migrant, and plan 504. All interactive reports, with the exception of the Longitudinal Data Report, allowed the user to provide a custom title for the report.

9.4.1 Roster Report

The Montana CRT *Roster Report* provides a roster of all students in a school and provides performance on the common items that are released to the public, one report per content area. For all grades and content areas, the student names and identification numbers are listed as row headers down the left side of the report. The items are listed as column headers in the same order they appeared in the released item document.

For each item, the following are shown:

- the depth of knowledge (DOK) code
- the item type
- the correct response key for multiple choice items
- the total possible points
- content standard

For each student, multiple-choice items are marked either with a plus sign (+), indicating that the student chose the correct multiple-choice response, or a letter (from A to D), indicating the incorrect response chosen by the student. For short-answer and constructed-response items, the number of points earned is shown. All responses to released items are shown in the report, regardless of the student's participation status. The columns on the right side of the report show the Total Test results, broken into several categories. Subcategory Points Earned columns show points earned by the student in each content area subcategory relative to total possible points. A Total Points Earned column is a summary of all points earned and total possible points in the content area. The last two columns show the student's scaled score and performance level. Students reported as Not Tested are given a code in the performance level column to indicate the reason the student did not test. It is important to note that not all items used to compute student scores are included in this report, only released items. At the bottom of the report, the average percentage correct for each multiple-choice item and average scores for the short-answer and constructed-response items are shown for the school, district, and state. When advanced filtering criteria are applied by the user, the School and District Percent Correct/Average Score rows at the bottom of the report are blanked out and only the Group row and the State row for the group selected will contain data. This report can be saved, printed, or exported as a PDF.

The Montana CRT roster is confidential and should be kept secure within the school and district. FERPA requires that access to individual student results be restricted to the student, the student's parents/guardians, and authorized school personnel.

9.4.2 Performance Level Summary

The Performance Level Summary provides a visual display of the percentages of students in each performance level for a selected grade/content area. The four performance levels (Novice, Nearing Proficiency, Proficient, and Advanced*)* are represented by various colors in a pie chart. A separate table is also included below the chart that shows the number and percentage of students in each performance level. This report can be saved, printed, or exported as a PDF or JPG file.

9.4.3 Item Analysis Data

The Released Items Summary Data report is a school-level report that provides a summary of student responses to the released items for a selected grade/content area. The report is divided into two sections by item type (multiple-choice and open-response). For multiple-choice items, the total number/percent of students who answered the item correctly and the number of students who chose each incorrect option or provided an invalid response are reported. An invalid response on a multiple-choice item is defined as "the item was left blank" or "the student selected more than one option for the item." For open-response items, point value and average score for the item are reported. Users are also able to view the actual released items within this report. If a user clicks on a particular magnifying glass icon next to a released item number, a pop-up box will open displaying the released item.

9.4.4 Longitudinal Data Report

The longitudinal data report is a confidential student-level report that provides individual student performance data for multiple test administrations. Results are reported for a student going back to academic year 2006–07. The state-assigned student identification number is used to link students across test administrations. Student performance on future test administrations will be included on this report over time. This report can be saved, printed, or exported as a PDF file.

9.5 Interpretive Materials and Workshops

An interpretive guide to the reports is provided on the OPI website: http://opi.mt.gov/.

9.6 Quality Assurance

Quality assurance measures are embedded throughout the entire process of analysis and reporting. The data processor, data analyst, and psychometrician assigned to work on Montana CRT implement quality control checks of their respective computer programs and intermediate products. Moreover, when data are handed off to different functions within the Data Services and Static Reporting (DSSR) and Psychometrics and Research (P&R) departments, the sending function verifies that the data are accurate before handoff. Additionally, when a function receives a data set, the first step is to verify the data for accuracy.

Another type of quality assurance measure is parallel processing. Different exclusions that determine whether each student receives scaled scores and/or is included in different levels of aggregation are parallel processed. Using the decision rules document, two data analysts independently write a computer program that assigns students' exclusions. For each content area and grade combination, the exclusions assigned by each data analyst are compared across all students. Only when 100% agreement is achieved can the rest of data analysis be completed.

Another level of quality assurance involves the procedures implemented by the quality assurance group to check the accuracy of reported data. Using a sample of schools and districts, the quality assurance group verifies that reported information is correct. The step is conducted in two parts: (1) verify that the computed information was obtained correctly through appropriate application of different decision rules, and (2) verify that the correct data points populate each cell in the Montana CRT reports. The selection of sample schools and districts for this purpose is very specific and can affect the success of the quality control efforts. There are two sets of samples selected that may not be mutually exclusive.

The first set includes those that satisfy the following criteria:

- One-school district
- Two-school district
- Multi-school district

The second set of samples includes districts or schools that have unique reporting situations as indicated by decision rules. This second set is necessary to ensure that each rule is applied correctly. The second set includes the following criteria:

- Private school
- School with excluded (not tested) students

The quality assurance group uses a checklist to implement its procedures. After the checklist is completed, sample reports are circulated for psychometric checks and program management review.

CHAPTER 10. VALIDITY

Because interpretations of test scores, and not a test itself, are evaluated for validity, the purpose of the *2009–10 Montana CRT Technical Report* is to describe several technical aspects of the Montana CRT tests in support of score interpretations (AERA, 1999). Each chapter contributes an important component in the investigation of score validation: test development and design; test administration; scoring, scaling, and equating; item analyses; reliability; and score reporting.

As stated in the overview chapter, *Standards for Educational and Psychological Testing* (AERA, et al., 1999) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. The evidence around test content, response processes, internal structure, relationship to other variables, and consequences of testing speak to different *aspects* of validity but are not distinct *types* of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

Evidence on test content validity is meant to determine how well the assessment tasks represent the curriculum and standards for each content area and grade level. Content validation is informed by the item development process, including how the test blueprints and test items align to the curriculum and standards. Viewed through this lens provided by the Standards, evidence based on test content was extensively described in Chapters 3 and 4. Item alignment with Montana content standards; item bias, sensitivity and content appropriateness review processes; adherence to the test blueprint; use of multiple item types; use of standardized administration procedures, with accommodated options for participation; and appropriate test administration training are all components of validity evidence based on test content. As discussed earlier, all CRT questions are aligned by Montana educators to specific Montana content standards, and undergo several rounds of review for content fidelity and appropriateness. Items are presented to students in multiple formats (constructed-response, short-answer and multiple-choice). Finally, tests are administered according to state-mandated standardized procedures, with allowable accommodations, and all test proctors are required to attend annual training sessions.

The scoring information in Chapter 5 describes the steps taken to train and monitor hand-scorers, as well as quality control procedures related to scanning and machine scoring. To speak to student response processes, however, additional studies would be helpful and might include an investigation of students' cognitive methods using think-aloud protocols.

Evidence based on internal structure is presented in great detail in the discussions of item analyses, reliability, and scaling and equating in Chapters 6 through 8. Technical characteristics of the internal structure of the assessments are presented in terms of classical item statistics (item difficulty, item-test correlation), differential item functioning analyses, dimensionality analyses, reliability, standard errors of measurement, and item response theory parameters and procedures. Each test is equated to the same grade and content test from the prior year in order to preserve the meaning of scores over time. In general, item difficulty and

discrimination indices were in acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that most items were assessing consistent constructs, and students who performed well on individual items tended to perform well overall.

Evidence based on the consequences of testing is addressed in the scaled scores information in Chapter 7 and the reporting information in Chapter 9, as well as in the test interpretation guide, which is a separate document that is referenced in the discussion of reporting. Each of these chapters speaks to the efforts undertaken to promote accurate and clear information provided to the public regarding test scores. Scaled scores offer the advantage of simplifying the reporting of results across content areas, grade levels, and subsequent years. Performance levels provide users with reference points for mastery at each grade level, which is another useful and simple way to interpret scores. Several different standard reports are provided to stakeholders. In addition, a data analysis tool is provided to each school system to allow educators the flexibility to customize reports for local needs. Additional evidence of the consequences of testing could be supplemented with broader investigation of the impact of testing on student learning.

To further support the validation of the assessment program, additional studies might be considered to provide evidence regarding the relationship of CRT results to other variables including the extent to which scores from the CRT converge with other measures of similar constructs, and the extent to which they diverge from measures of different constructs. Relationships among measures of the same or similar constructs can sharpen the meaning of scores and appropriate interpretations by refining the definition of the construct.

The evidence presented in this report supports inferences of student achievement on the content represented on the Montana content standards for reading, mathematics, and science for the purposes of program and instructional improvement and as a component of school accountability.

REFERENCES

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory.* Belmont, CA: Wadsworth, Inc.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Baker, F. B., & Kim, S-H. (2004). *Item response theory: parameter estimation techniques* (2nd ed.). New York: Marcel Dekker, Inc.

Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). Fort Worth: Holt, Rinehart and Winston.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, *23*, 355–368.

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description. In P. W. Holland & H. Wainer (Eds.) *Differential item functioning* (pp. 35–66). Hillsdale, NJ: Lawrence Erlbaum Associates.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York: John Wiley and Sons, Inc.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer Academic Publishers.

Hambleton, R. K., Swaminathan, H., & Rogers, J. H. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, Inc.

Hambleton, R. K., & van der Linden, W. J. (1997). *Handbook of modern item response theory*. New York, NY: Springer-Verlag.

Joint Committee on Testing Practices (2004). *Code of fair testing practices in education*. Washington, DC: Joint Committee on Testing Practices. Available from www.apa.org/science/programs/testing/fair-code.aspx

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, *32*, 179–197.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Muraki, E., & Bock, R. D. (2003). PARSCALE 4.1. Lincolnwood, IL: Scientific Software International.

Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 221–262).

Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika, 52,* 589–617.

Stout, W. F., Froelich, A. G., & Gao, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. In A. Boomsma, M. A. J. van Duign, & T. A. B. Snijders (Eds.), *Essays on item response theory,* (pp. 357–375). New York: Springer-Verlag.

Zhang, J., & Stout, W. F. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika, 64,* 213–249.

APPENDICES

Appendix A—Analysis and Reporting Decision Rules

Analysis and Reporting Decision Rules
Montana Comprehensive Assessment System (MontCAS) CRT and CRT-Alternate (Final)
Spring 09-10 Administration

This document details rules for analysis and reporting. The final student level data set used for analysis and reporting is described in the "Data Processing Specifications." This document is considered a draft until the Montana Office of Public Instruction (OPI) signs off. If there are rules that need to be added or modified after said sign-off, OPI sign off will be obtained for each rule. Details of these additions and modifications will be in the Addendum section.

I. General Information

A. Tests Administered

Grade	Subject	Items included in Raw Score		IABS Reporting Categories (Standards) (Not Applicable for CRT-Alternate)
		CRT	CRT-Alt	
03	Reading Math	Common	All	Cat2
04	Reading Math	Common	All	Cat2
	Science	Common	All	Cat3
05	Reading Math	Common	All	Cat2
06	Reading Math	Common	All	Cat2
07	Reading Math	Common	All	Cat2
08	Reading Math	Common	All	Cat2
	Science	Common	All	Cat3
10	Reading Math	Common	All	Cat2
	Science	Common	All	Cat3

B. Reports Produced
1. Student Labels (Printed)
2. Student Report (Printed and posted online)
3. Roster & Item Level Report (CRT-Alt: posted online; CRT:Interactive System)
   - by grade, subject and class/group
4. Summary Report (Online)
   Consists of sections:
   I. Distribution of Scores
   II. Subtest Results
   III. Results for Subgroups of Students
   - by grade, subject and school

- by grade, subject and system
- by grade, subject (state level)

C. Files Produced (Format: comma delimited format)
   1. One state file for each grade
      a. Consists of student level results
      b. Alternately assessed students are in separate files by grade.
      c. Naming conventions
         i. CRT All subjects- Studentdatafile[2 digit grade].csv
         ii. CRT-Alternate All subjects- altStudentdatafile[2 digit grade].csv
      d. File layout: Studentdatafilelayout.xls and altstudentdatafilelayout.xls

   2. System level files (Format: Excel ; Online)
      a. Consists of student level results for each system for each grade. Contains all subjects tested at that grade.
      b. Naming convention: Studentdatafile[2 digit grade].xls
      c. File Layout: Systemstudentdatafilelayout.xls

   3. School level file (Format: Excel; Online)
      a. Consists of previous year's student level results for each school and grade. Contains all subjects tested at that grade.
      b. Naming convention: Studentdatafile[2 digit grade].xls
      c. File Layout: Systemstudentdatafilelayout.xls

D. School Type

Schtype	**Source**	**Description**	**Included in Aggregations**		
			School	**System**	**State**
"Pras"	Data file provided by state	Private Accredited School. They are their own system	Yes. Same information for school & system but both sets of reports produced	Yes. Same information for school & system but both sets of reports produced	No
"Prnas"	Data file provided by state	Private non-accredited school. They are their own system	Yes. Same information for school & system but both sets of reports produced	Yes. Same information for school & system but both sets of reports produced	No
"SNE"	Scanned data/updated by OPI	Student not enrolled	No.	No.	No.
"Oth"		non-private school	Yes	Yes	Yes

E. Other Information
   1. CRT Tests are constructed with a combination of common and embedded field test items.
   2. The CRT-Alternate consists of a set of 5 performance tasklets. The number of items in each tasklet varies.
   3. Braille Students:
      a. See Appendix A.1 for a list of the items not included in the Braille form.
      b. If a student is identified as taking the Braille test, these items are not included in the student's raw score. The student is scaled on a separate form based on the items that are available to him or her. See the Calculations section for more information.

II. Student Participation/Exclusions

A. Test Attempt Rules
   1. A valid response to a multiple choice item is A, B, C, or D. An asterisk (multiple marks) is not considered a valid response. A valid score for an open response item is a non-blank score.
   2. Incomplete (CRT): The student has fewer than two (2) but at least one (1) valid response to common items.
   3. Incomplete (CRT-Alternate): The student has fewer than three (3) scores across all tasklets.
   4. The student is classified as Did Not Participate (DNP) in CRT if the student does not have any valid responses for that subject in either CRT or CRT-Alternate and no not tested reason.

B. Not Tested Reasons
   1. If a student is marked First year LEP regardless of items attempted the student is considered first year LEP for reporting purposes. Reading is optional for first year in U.S schools LEP students.

C. Student Participation Status
   1. The following students are excluded from all aggregations.
      a. Foreign Exchange Students (FXS).
      b. Homeschooled students (schtype='SNE').
      c. Student in school less than 180 hours (PSNE).
      d. DNP (for that subject)
      e. First year in U.S schools LEP*(regardless of how many items were attempted)
      f. CRT only: Student tested with Non-Standard Accommodations (NSA for that subject)*
      g. Alt (Alt='1')

      * These students are aggregated on the Disaggregated report in their respective rows.
   2. If any of the non-standard accommodations are bubbled the student is considered tested with non-standard accommodations (NSA) in that subject.
   3. If the student has not been in that school for the entire academic year the student is excluded from school level aggregations (NSAY).
   4. If the student has not been in that system for the entire academic year the student is excluded from system and school level aggregations (NDAY).
   5. If the student took the alternate assessment the student is not counted as participating in the general assessment. Alternate Assessment students receive their results on an Alternate Assessment Student Report. They are reported according to participation rules stated in this document.

6. (CRT-Alternate) If the teacher halted the administration of the assessment after the student scored zero (0) for three (3) consecutive items within tasklets , the student is classified as Halted in that subject. If the student was halted within a tasklet then the rest of the items within the tasklet are blanked out and do not count toward the student's score. If the other tasklets are complete then those items will be counted toward the student's score.
7. If the student took the Braille form of the test the raw scores are not included in raw score school, system or state averages. They are not included in group averages on the interactive roster.

D. Student Participation Summary

Participation Status	**Part. Flag**	**Raw score**	**Scaled Score**	**Perf. level**	**Included on Roster**	**Included in aggregations**		
						Sch	**Sys**	**Sta**
FXS	E	✓	✓	✓				
SNE	E	✓	✓	✓				
PSNE	E	✓	✓	✓				
NSA(by subject) Applies to CRT only	A	✓	✓	✓	✓	Only included in count and percents on Disaggregated report for nonstandard accommodations.		
First year in U.S schools LEP	A	✓	See Report Specific Rules	See Report Specific Rules	✓			
NSAY only	B	✓	✓	✓	✓		✓	✓
NDAY	C	✓	✓	✓	✓			✓
ALT*	A	✓	✓	✓	✓	See footnote below		
Incomplete	A	✓	✓	✓	✓			
DNP (Non-Participants)	F	✓	✓	✓	✓			
Halted(CRT-Alt only by subject)	D	✓	✓	✓	✓	✓	✓	✓
Tested	Z	✓	✓	✓	✓	✓	✓	✓

* They are included in summary data only for alternate assessment reports (according to participation rules).

If a student has conflicting participation statuses the following hierarchy is applied to determine how the student is reported:

F (Student attempted no items and is not alt and cannot be classified as first-year LEP)
E (FXS, SNE or PSNE)
A (NSA, first year in U.S schools LEP, ALT or INC)
C (NDAY)

B (NSAY)
D (Halted; applies to CRT-Alt only)
Z (completed CRT or CRT-Alt and none of the above conditions apply)

III. Calculations

A. Raw Scores
   1. (CRT) Raw scores are calculated using the scores on common multiple choice and open response items.
   2. (CRT-Alternate) Raw score is the sum of the individual item scores.

B. Scaling
   1. Scaling is accomplished by defining the unique set of test forms for each grade/subject combination. This is accomplished as follows:
      a. Translate each form and position into the unique item number assigned to the form/position.
      b. Order the items by
         I. Type- multiple choice, short-answer, constructed-response
         II. Form-common, then by ascending form number.
         III. Position
      c. If an item number is on a form, then set the value for that item number to '1', otherwise set to '.'. Set the exception field to '0' to indicate this is an original test form.
      d. If an item number contains an 'X' (item is not included in scaling) then set the item number to '.'. Set the exception field to '1' to indicate this is not an original test form.
      e. Compress all of the item numbers together into one field in the order defined in step II to create the test for the student.
      f. Select the distinct set of tests from the student data and order by the exception field and the descending test field.
      g. Check to see if the test has already been assigned a scale form by looking in the daScaleForm table. If the test exists then assign the existing scale form. Otherwise assign the next available scale form number. All scale form numbering starts at 01 and increments by 1 up to 99.

   2. Psychometrics provides a lookup table for each scale form. These lookup tables are used to assign scaled scores, performance levels and standard errors.
   3. The scaled score cuts for all three subjects and all grades have been fixed and are the same as last year for the CRT.
   4. Students excluded from aggregations at the state level are excluded from psychometric files.

C. CRT-Alternate: The classcode is created using the following steps:
   1. The following students are not included when creating the class codes.
      - SNE
      - FXS
      - PSNE
   2. The dataset (by grade) is sorted by schcode and class/group name
   3. The records are then numbered consecutively starting at 1. This number is then padded with zeros (in front) to create a 3 digit number.

D. Performance Level coding:

Numeric Performance Level	Performance level Name	Abbreviation
1(lowest)	Novice	N
2	Nearing Proficient	NP
3	Proficient	P
4(highest)	Advanced	A

E. Rounding Table

Calculation	Rounded (to the nearest)
Static Reports: Percents and averages	Whole number
Item averages : Multiple choice items	The average is multiplied by 100 and rounded to the nearest whole number.
Item averages: Open response items	Open-response item averages are rounded to the nearest tenth.

F. Minimum N size
1. The number of included students (N) in a subject is the number of students in the school/system/state minus FXS minus PRAS minus PRNAS minus PSNE minus SNE minus First year LEP minus Incomplete minus NSA minus DNP.
2. Minimum N size is 10.
3. School/system reports are produced regardless of N-size.

G. The common items are used in reporting the average number of points for each standard.

H. Assignment of rperflevel
1. If the student is marked as taking the CRT-Alt then rperflevel='A' otherwise
2. If the student is classified as did not participate (DNP) then rperflevel='D' otherwise
3. If the student is Incomplete in a subject and not marked first year LEP rperflevel='I' otherwise
4. If the student is incomplete in Reading or has not attempted any items in Reading and is marked first year LEP rperflevel='L' for all subjects otherwise
5. If the student does not meet any of the above conditions then rperflevel=perflevel.

IV. Report Specific Rules

A. Student Label
1. If a student is First year LEP and incomplete in Reading, the Reading performance level is 'LEP'. The reading scaled score is blank.

2. If a student is First year LEP, the math and science performance levels are the name of the earned performance level and the scaled scores are the student's earned score.
3. If the student is not first year LEP, the performance level name corresponding to the student's earned score is displayed.
4. If the student is First year LEP but is not incomplete in Reading then the student receives his earned scaled score and performance level.
5. If the student is DNP the student receives a student label. The student receives scaled score =200 and performance level=Novice.
6. The student's name is formatted as Lname, Fname.
7. The student's name is uppercase.
8. The school and system names are title case.
9. The labels are sorted alphabetically by Lname, Fname within school and grade.
10. Test date is 2010.
11. Performance level name from section III.D above is shown on the label if the student receives a performance level.

B. Roster & Item Level Report-Alternate Assessment only
   1. If a student is First year LEP and the student is not incomplete in Reading:
      a. The math (and science) performance level is the abbreviation of the earned performance level and the scaled score is the student's earned score.
      b. The reading performance level is the abbreviation of the earned performance level and the scaled score is the student's earned score.
      c. The student is excluded from Reading, Math and Science aggregations.
   2. If the student is First year LEP and incomplete in Reading
      a. The student's Reading, Math (and Science) performance levels are 'LEP'.
      b. The student's math (and science) scaled score is the student's earned scaled score and the reading scaled score is blank.
      c. The student's responses for all subjects are displayed.
      d. The student is excluded from Math, Reading (and Science) aggregations.
   3. If the student is not first year LEP, the performance level abbreviation corresponding to the student's earned score is displayed.
   4. If the student is incomplete the student receives the scores with a footnote (†) "Student did not complete the assessment."
   5. There is no last name or first name for the student, the name displayed is "Name Not Provided". These students appear at the bottom of the roster.
   6. If class/group information is missing the roster is done at the school level.
   7. Results for Alternate Assessment students are reported only on their class/group/school's alternate *Roster & Item Level Report.*
   8. Within each demonstration school the class is 'DEM'.
   9. Only the standards reported on the Summary report are reported on the roster.
   10. The student's are sorted by lname, fname
   11. Student names are formatted Lname, Fname.
   12. Student names are uppercase.
   13. Performance level abbreviation from section III.D is placed the performance level column if the student receives are performance level.
   14. If NSAY='1' or NDAY='1' then place appropriate symbol beside the first name. See addenda section for symbols
   15. If [subject]halted='1' for any subject then place appropriate symbol beside the first name. See addenda section for symbols.

C. Interactive Roster – CRT only

1. Students who are DNP in a subject are reported with scaled score=200 and performance level='DNP'.
2. Students who are Incomplete in a subject are reported with their earned scaled score and performance level='INC' on the interactive roster.
3. Students who are first-year LEP and who complete the reading test are reported with their earned scaled score and performance level and are included in school, system and state level aggregations for all subjects unless otherwise excluded based on completeness in math or science.
4. Students who are first-year LEP and who do not complete the reading test are reported with their earned scaled score and performance level='LEP' for all subjects. These students are excluded from school, system and state level aggregations.
5. Students who participated in Alternate assessment are listed on the rosters. Their scaled score is blank and the performance level='ALT'. These students are not included in aggregations.
6. The items are reported using the released item number.
7. Students who took the Braille form are not included in any rawscore aggregations. These students have a scaleform other than 01.
8. The following students will have included set to 0 in tblscoreditem (these students are excluded from performance level aggregations):
    a. The student did not participate in the subject (partstatus='F')
    b. The student has partstatus='E'
    c. The student is LEPfirst (LEPfirst='1' regardless of how many items attempted)
    d. The student is incomplete in the subject.
    e. The student took the alternate assessment (alt='1')
    f. Student took the subject with nonstandard accommodations (NSA).
    g. Student is NSAY (NSAY='1').
    h. Student is NDAY (NDAY='1').
9. If the student took the Braille form (Braille='1'), included is set to 2. These students are excluded from raw score aggregations.
10. If students do not fall into any of the categories in numbers 8 and 9 above, included is set to '1'.
11. If partstatus='E' for any subject then interactive='0' otherwise interactive='1'. Students with interactive='0' are not available in the interactive site.
12. State level item averages do not include students with school type PRAS, PRNAS or SNE.
13. District level item averages do not include students who are marked nday='1'.
14. Only students whose partstatus is not 'E" for any subject are included in tblStuLongitudinal.
15. The filter column in tblItemAveragesLookup is the concatenation of the gender,ethnic,iep,lep,econdis,migrant and plan504 fields in that order. The leading zero in the ethnic field is removed prior to concatenating.
16. RepType='0' for all records in tblItemAverages.

D. Summary Report

1. Section III (Results for Subgroups of Students)
    a. Performance level results for subgroups with N less than 10 are suppressed. N is always reported. Footnote * 'Less than 10 students were assessed.'
    b. CRT only: Count of students who are considered NSA for that subject excluding those students who are incomplete, nsay (at school level), nday (at school and

system level) or FXS or SNE or PSNE or First year LEP or alt (general assessment report).

c. Count of First year LEP students excludes those students who are nsay (at school level), nday (at school or system level) or incomplete or FXS or SNE or PSNE or NSA or alt (general assessment).

2. Section II (Subtest Results) Students with scaleform other than 01 are not included in Subtest Results.

V. Data File Rules

1. The following students are not included in the state file:
   a. Alternate Assessment students (in CRT)
   b. Homeschooled students (SNE)
   c. Student is in school less than 180 hours (PSNE)
2. If the student receives a performance level 'LEP' on the student report in Reading, the student receives LEP for the Reading performance level in the state files.
3. Alt students who are halted are marked '1' in the halted field for that subject.
4. Students who take the Braille form of the test are flagged Braille='1' in the state and system level files.
5. In the system level files only the released scored items are included.
6. The following students are not included in the system level files:
   a. Foreign Exchange students (FXS='1')
   b. Homeschooled students (SNE)
   c. Student is in school less than 180 hours (PSNE)
7. The following students are not included in the previous year school level files:
   a. Foreign Exchange students (FXS='1')
   b. Homeschooled students (SNE)
   c. Student is in school less than 180 hours (PSNE)

VI. Shipping Product Code Summary

1. School (ReportFor='1')

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
03	Student Labels (CRT)	03	Reading and Math	00	1 set for each school
04	Student Labels (CRT)	03	Reading, Math and Science	00	1 set for each school
05	Student Labels (CRT)	03	Reading and Math	00	1 set for each school
06	Student Labels (CRT)	03	Reading and Math	00	1 set for each school

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
07	Student Labels (CRT)	03	Reading and Math	00	1 set for each school
08	Student Labels (CRT)	03	Reading Math and Science	00	1 set for each school
10	Student Labels (CRT)	03	Reading Math and Science	00	1 set for each school
03	Student Report (CRT)	02	Reading and Math	00	1 for each student
04	Student Report (CRT)	02	Reading Math and Science	00	1 for each student
05	Student Report (CRT)	02	Reading Math	00	1 for each student
06	Student Report (CRT)	02	Reading and Math	00	1 for each student
07	Student Report (CRT)	02	Reading and Math	00	1 for each student
08	Student Report (CRT)	02	Reading Math and Science	00	1 for each student
10	Student Report (CRT)	02	Reading Math and Science	00	1 for each student
03	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
04	Student Labels (CRT-Alt)	07	Reading, Math and Science	00	1 set for each school
05	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school
06	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school
07	Student Labels (CRT-Alt)	07	Reading and Math	00	1 set for each school
08	Student Labels (CRT-Alt)	07	Reading Math and Science	00	1 set for each school
10	Student Labels (CRT-Alt)	07	Reading Math and Science	00	1 set for each school
03	Student Report (CRT-Alt)	08	Reading and Math	00	1 for each student
04	Student Report (CRT-Alt)	08	Reading Math and Science	00	1 for each student
05	Student Report (CRT-Alt)	08	Reading Math	00	1 for each student
06	Student Report (CRT-Alt)	08	Reading and Math	00	1 for each student

Grade	Report Name	ReportType	Subject	ContentCode	Quantity
07	Student Report (CRT-Alt)	08	Reading and Math	00	1 for each student
08	Student Report (CRT-Alt)	08	Reading Math and Science	00	1 for each student
10	Student Report (CRT-Alt)	08	Reading Math and Science	00	1 for each student
00	Interp. Guide	04		00	1 per school

Addenda:

A. PDF file naming conventions to be used by Report Programmer
   1. Printed Reports
      a. Labels
         MT La [grade].pdf
      b. Student Report (Parent Copy)
         #####[systemcode]MT Sr [grade] (Parent Copy).pdf
      c. Student Report (School Copy)
         #####[systemcode]MT Sr [grade] (School Copy).pdf

   2. Web Reports
      a. School Summary Reports
         MT Su Sch [3 character subject][grade].pdf
      b. System Summary Reports
         MT Su Dis [3 character subject][grade].pdf
      c. State Summary Reports
         MT Su Sta [3 character subject][grade].pdf

B. Footnotes to be placed on the bottom of the roster. These footnotes should be on all pages for all rosters.
   † Student did not complete the assessment.
   ¥ Not in school and/or system for full academic year.
   § Teacher halted the administration of one or more of the five tasklets after the student scored a 0 for three consecutive items within a tasklet on two different test administrations. Any completed tasklets have been scored and are reflected in the student's scaled score.

C. Section III.H Assignment of rperflevel applies only to CRT.

Appendix A

1. Items not available on the Braille form

Grade	Subject	Form	Position	Reporting Category
03	Mat	00	25	2
03	Mat	01	73	
04	Rea	00	21	2
04	Mat	00	10	4
04	Mat	00	45	2
04	Mat	01	50	
04	Sci	00	23	3
04	Sci	01	65	
10	Rea	00	60	2
10	Mat	00	23	6
10	Mat	01	32	
10	Mat	00	52	4
10	Sci	00	21	3
10	Sci	00	23	1

Appendix B—Accommodation Frequencies by Content Area

Table B-1. 2009–10 MontCAS: Numbers of Students Tested With Accommodations by Accommodation Type and Grade – Mathematics

Accommodation code	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Grade 7*	*Grade 8*	*Grade 10*
matAccom01	191	171	199	134	156	108	134
matAccom02	363	384	382	354	348	288	205
matAccom04	263	246	207	143	106	90	71
matAccom05	1303	1432	1250	1093	778	705	571
matAccom06	211	194	179	127	134	111	65
matAccom07	976	942	867	645	366	361	277
matAccom08	1026	1040	865	696	425	307	290
matAccom09	8	2	2	2	1	2	3
matAccom10	6	1	4	6	32	49	9
matAccom12	1	2	1	1	4	1	0
matAccom13	3	4	4	5	0	1	4
matAccom14	9	6	11	6	0	1	0
matAccom15	3	1	5	2	1	0	0
matAccom16	0	0	3	5	8	9	1
matAccom17	0	0	0	0	0	0	1
matAccom18	4	1	4	7	5	3	2
matAccom19	126	143	118	66	66	64	10
matAccom20	12	12	6	6	17	7	3
matAccom21	0	0	2	10	0	2	0
matAccom22	1287	1294	1055	777	583	517	217
matAccom23	7	10	3	8	2	5	20
matAccom24	43	77	85	63	24	21	16
matAccom25	111	85	133	101	89	101	60
matAccom26	2	1	0	1	0	0	1
matAccom27	4	3	6	6	8	5	2
matAccom28	1	0	0	1	2	0	1
matAccom30	0	0	0	0	0	0	0
matAccom32	0	0	0	0	0	0	0

Table B-2. 2009–10 MontCAS: Accommodations – Mathematics

Accommodation	*Description*
matAccom01	Change in Administration Time: Test is administered at a time of day or a day of the week based on student needs.
matAccom02	Session Duration: Test is administered in appropriate blocks of time for individual student needs, followed by rest breaks.
matAccom04	Individual Administration: Test is administered in a one-to-one situation.
matAccom05	Small Group Administration: Test is administered to a small group of students.
matAccom06	Reduce Distracters: Student is seated at a carrel or other physical arrangement that reduces visual distractions.
matAccom07	Alternative Setting: Test is administered to a student in a different setting.
matAccom08	Change in Personnel: Test is administered by other personnel known to the student (e.g., LEP, Title I, special education teacher).
matAccom09	Home Setting: Test is administered to the student by school personnel in their home.
matAccom10	Front Row Seating: Student is seated at the front of the classroom when taking the test.
matAccom12	Magnification: Student used equipment to magnify test materials.
matAccom13	Student (not groups of students) wears equipment to reduce environmental noises.
matAccom14	Template: Student uses a template. An example is a piece of card stock that has a window cut out that enables the student to focus by isolating lines of text or items.
matAccom15	Amplification: Student uses amplification equipment (e.g., hearing aid or auditory trainer) while taking test.
matAccom16	Writing Tools: Student uses a typewriter or word processor (without activating spell check).
matAccom17	Voice Activation: Student speaks response into computer equipped with voice-activation software.
matAccom18	Bilingual Dictionary: Student uses a bilingual dictionary.
matAccom19	Dictation: Student dictates answers to a test administrator who records them in the Answer Booklet.
matAccom20	Writing Tools: Student marks or writes answers with the assistance of a technology device or special equipment.
matAccom21	Assistive Technology: Another form of assistive technology routinely used by the student (that does not change intent or test content).
matAccom22	Oral Presentation: The test administrator must read the test items and answer choices word-for-word. Before reading aloud, the test administrator should advise students that each item and answer choice will be read aloud in exactly the order as presented. Students should also be advised that items, including answer choices, will be repeated at the end of a session in case the students wish to review/check their work.
matAccom23	Test Interpretation: Tests, including directions, are interpreted for students who are deaf or hearing-impaired.
matAccom24	Test Directions with Verification: An administrator gives test directions with verification (by using a highlighter) so that student understands them.
matAccom25	Test Directions Support: An administrator assists student in understanding test directions, including giving directions in native language.
matAccom26	Braille: Braille version of the test was used by the student.
matAccom27	Large Print: A large-print version of the test is used by student.
matAccom28	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.
matAccom30	Student uses a calculator, number chart, arithmetic table, or manipulative on the no calculator sections of the mathematics test.
matAccom32	With verification from the OPI in advance of the testing window, some other approved accommodation is used by the student.

Table B-3. 2009–10 MontCAS: Numbers of Students Tested With Accommodations by Accommodation Type and Grade – Reading

Accommodation code	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Grade 7*	*Grade 8*	*Grade 10*
REAAccom01	198	167	192	131	157	105	144
REAAccom02	355	377	394	353	360	295	224
REAAccom04	250	249	188	136	104	94	77
REAAccom05	1277	1374	1205	1087	760	696	589
REAAccom06	207	191	162	120	140	118	71
REAAccom07	948	895	853	630	368	376	292
REAAccom08	1002	1007	846	687	422	319	310
REAAccom09	9	2	2	2	1	2	4
REAAccom10	7	1	4	6	35	49	9
REAAccom12	1	3	3	1	4	1	0
REAAccom13	3	3	4	6	0	2	3
REAAccom14	10	7	13	6	0	2	0
REAAccom15	3	1	5	2	1	1	0
REAAccom16	0	0	6	7	11	18	11
REAAccom17	0	0	0	0	0	0	1
REAAccom18	4	3	4	8	3	3	2
REAAccom19	193	178	159	102	85	91	15
REAAccom20	14	12	5	7	20	8	1
REAAccom21	0	0	1	2	0	0	0
REAAccom22	1093	1017	836	642	470	468	215
REAAccom23	7	10	1	4	2	6	20
REAAccom24	43	77	81	66	23	21	22
REAAccom25	102	79	110	98	98	104	57
REAAccom26	0	1	0	1	1	1	1
REAAccom27	4	2	6	5	8	5	2
REAAccom28	2	1	0	0	2	0	1
REAAccom29	0	0	0	0	0	0	0
REAAccom31	0	0	0	0	0	0	0

Table B-4. 2009–10 MontCAS: Accommodations – Reading

Accommodation	*Description*
REAAccom01	Change in Administration Time: Test is administered at a time of day or a day of the week based on student needs.
REAAccom02	Session Duration: Test is administered in appropriate blocks of time for individual student needs, followed by rest breaks.
REAAccom04	Individual Administration: Test is administered in a one-to-one situation.
REAAccom05	Small Group Administration: Test is administered to a small group of students.
REAAccom06	Reduce Distracters: Student is seated at a carrel or other physical arrangement that reduces visual distractions.
REAAccom07	Alternative Setting: Test is administered to a student in a different setting.
REAAccom08	Change in Personnel: Test is administered by other personnel known to the student (e.g., LEP, Title I, special education teacher).
REAAccom09	Home Setting: Test is administered to the student by school personnel in their home.
REAAccom10	Front Row Seating: Student is seated at the front of the classroom when taking the test.
REAAccom12	Magnification: Student used equipment to magnify test materials.
REAAccom13	Student (not groups of students) wears equipment to reduce environmental noises.
REAAccom14	Template: Student uses a template. An example is a piece of card stock that has a window cut out that enables the student to focus by isolating lines of text or items.
REAAccom15	Amplification: Student uses amplification equipment (e.g., hearing aid or auditory trainer) while taking test.
REAAccom16	Writing Tools: Student uses a typewriter or word processor (without activating spell check).
REAAccom17	Voice Activation: Student speaks response into computer equipped with voice-activation software.
REAAccom18	Bilingual Dictionary: Student uses a bilingual dictionary.
REAAccom19	Dictation: Student dictates answers to a test administrator who records them in the Answer Booklet.
REAAccom20	Writing Tools: Student marks or writes answers with the assistance of a technology device or special equipment.
REAAccom21	Assistive Technology: Another form of assistive technology routinely used by the student (that does not change intent or test content).
REAAccom22	Oral Presentation: Only the questions and answer choices may be read aloud to the student. It is advised that the questions be read aloud to the student before he/she reads each passage. After the student has read the passage, the test administrator must read the questions and answer choices word-for-word one at a time in exactly the order as presented.
REAAccom23	Test Interpretation: Tests, including directions, are interpreted for students who are deaf or hearing-impaired.
REAAccom24	Test Directions with Verification: An administrator gives test directions with verification (by using a highlighter) so that student understands them.
REAAccom25	Test Directions Support: An administrator assists student in understanding test directions, including giving directions in native language.
REAAccom26	Braille: Braille version of the test was used by the student.
REAAccom27	Large Print: A large-print version of the test is used by student.
REAAccom28	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.
REAAccom29	Reading passages are read aloud to student, or student uses text-reader software for reading passages.
REAAccom31	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.

Table B-5. 2009–10 MontCAS: Numbers of Students Tested With Accommodations by Accommodation Type and Grade – Science

Accommodation code	*Grade 4*	*Grade 8*	*Grade 10*
sciAccom01	163	113	142
sciAccom02	361	261	229
sciAccom04	236	110	73
sciAccom05	1151	660	603
sciAccom06	192	113	74
sciAccom07	835	366	297
sciAccom08	948	296	309
sciAccom09	2	2	3
sciAccom10	1	48	9
sciAccom12	2	1	0
sciAccom13	4	2	2
sciAccom14	6	1	0
sciAccom15	1	0	0
sciAccom16	0	11	2
sciAccom17	0	0	1
sciAccom18	1	3	2
sciAccom19	164	81	13
sciAccom20	12	9	3
sciAccom21	0	1	0
sciAccom22	1139	494	247
sciAccom23	9	5	24
sciAccom24	73	22	21
sciAccom25	84	85	56
sciAccom26	1	0	1
sciAccom27	2	5	2
sciAccom28	0	0	1
sciAccom33	0	0	0

Table B-6. 2009–10 MontCAS: Accommodations – Science

Accommodation	*Description*
sciAccom01	Change in Administration Time: Test is administered at a time of day or a day of the week based on student needs.
sciAccom02	Session Duration: Test is administered in appropriate blocks of time for individual student needs, followed by rest breaks.
sciAccom04	Individual Administration: Test is administered in a one-to-one situation.
sciAccom05	Small Group Administration: Test is administered to a small group of students.
sciAccom06	Reduce Distracters: Student is seated at a carrel or other physical arrangement that reduces visual distractions.
sciAccom07	Alternative Setting: Test is administered to a student in a different setting.
sciAccom08	Change in Personnel: Test is administered by other personnel known to the student (e.g., LEP, Title I, special education teacher).
sciAccom09	Home Setting: Test is administered to the student by school personnel in their home.
sciAccom10	Front Row Seating: Student is seated at the front of the classroom when taking the test.
sciAccom12	Magnification: Student used equipment to magnify test materials.
sciAccom13	Student (not groups of students) wears equipment to reduce environmental noises.
sciAccom14	Template: Student uses a template. An example is a piece of card stock that has a window cut out that enables the student to focus by isolating lines of text or items.
sciAccom15	Amplification: Student uses amplification equipment (e.g., hearing aid or auditory trainer) while taking test.
sciAccom16	Writing Tools: Student uses a typewriter or word processor (without activating spell check).
sciAccom17	Voice Activation: Student speaks response into computer equipped with voice-activation software.
sciAccom18	Bilingual Dictionary: Student uses a bilingual dictionary.
sciAccom19	Dictation: Student dictates answers to a test administrator who records them in the Answer Booklet.
sciAccom20	Writing Tools: Student marks or writes answers with the assistance of a technology device or special equipment.
sciAccom21	Assistive Technology: Another form of assistive technology routinely used by the student (that does not change intent or test content).
sciAccom22	Oral Presentation: The test administrator must read the test items and answer choices word-for-word and in exactly the order as presented.
sciAccom23	Test Interpretation: Tests, including directions, are interpreted for students who are deaf or hearing-impaired.
sciAccom24	Test Directions with Verification: An administrator gives test directions with verification (by using a highlighter) so that student understands them.
sciAccom25	Test Directions Support: An administrator assists student in understanding test directions, including giving directions in native language.
sciAccom26	Braille: Braille version of the test was used by the student.
sciAccom27	Large Print: A large-print version of the test is used by student.
sciAccom28	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.
sciAccom33	Other: With verification from the OPI in advance of the testing window, some other approved accommodation is used by student.

Appendix C—Item Review Committee Members

PASSAGE REVIEW COMMITTEE MEMBERS

December 3-4, 2008

First name	*Last name*	*Position*
Richard	Desch	Test Coordinator
Connie	Filesteel	Instructor
Keith	Grebetz	Reading Teacher
Linda	Jones	Reading Teacher
Shelly	Moen	Teacher
Vicky	Panasuk	Title I Teacher
Cory	Pierce	Technology/Dean of Students
Lavada	Pilling	Reading Instructor
Penny	Reynolds	Title I Teacher
Carol	Shipley	Sped Teacher
Violet	Sinclair	Test Coordinator
Mona	Sindelar	Test Coordinator
Corri	Smith	Indian Ed Instructor

BENCHMARKING COMMITTEE MEMBERS

May 4-7, 2009

First name	*Last name*	*Position*
Beckie	Frisbee	Math
Bette	Paskey	Math
Cindy	Gremaux	Math
Rodd	Zeiler	Math
Paul	Tackes	Science
Nina	Miller	Science
Linda	Jones	Reading
Vicky	Panasuk	Reading

BIAS REVIEW COMMITTEE MEMBERS

May 11, 2009

First name	*Last name*	*Position*
Richard	Desch	Curriculum Coordinator
Pam	Diamond	8th Grade Language Arts Teacher
Chuck	Gameon	Elementary Principal
Cynthia	Green	Teacher
Callie	Kolste	Supervising Teacher/Teacher K-3
Ann	Magee	Teacher
Laura	Monroe	Special Education
Carol	Morgan	5th Grade Teacher
Mona	Sindelar	District Testing Coordinator

ITEM REVIEW COMMITTEE MEMBERS

May 12-13, 2009

First name	*Last name*	*Position*
Jonna	Ascheman	Reading 3/4
Marilyn	Beers	Reading 10
Lee	Brown	Mathematics 10
Katie	Burke	Science 4
Carl	Christiansen	Mathematics 7/8
Gale	Decker	Mathematics 5/6
Ashley	Gillespie	Science 8
Keith	Grebetz	Reading 5/6
Angela	Haas	Science 10
Judy	Haefner	Mathematics 5/6
Sheila	Hall	Reading 3/4
Amy	Hammill	Mathematics 10
Heidi	Hanks	Mathematics 5/6
Karen	Hough	Mathematics 3/4
Jean	Howard	Mathematics 7/8
Mike	Jetty	Science 8
Linda	Jones	Reading 7/8
Sara	Keast	Reading 10
Sheryl	Kohl	Mathematics 7/8
Callie	Kolste	Reading 3/4
Michael	Lanier	Mathematics 10
Mary	Lyndes	Mathematics 3/4
Jay	Mcgraw	Mathematics 7/8
Karen	Miller	Science 4
Michael	Munson-lenz	Science 4
Shannon	Murphy	Reading 5/6
Vicky	Panasuk	Reading 3/4
Christine	Perkins	Reading 7/8
Janice	Petritz	Science 8
Cory	Pierce	Mathematics 3/4
Roberta	Ray Kipp	Reading 7/8
Lynda	Reese	Reading 5/6
Mona	Shortman	Mathematics 3/4
Violet	Sinclair	Reading 7/8
Lydia	Tande	Mathematics 5/6
Adam	Wagner	Science 10
Debra	Westrom	Mathematics 5/6

OPERATIONAL TEST REVIEW COMMITTEE MEMBERS

July 21-22, 2009

First name	*Last name*	*Position*
David	Bixby	5th Grade Teacher
Tim	Bolten	8th Grade Math/Science Teacher
Marie	Boothe	High School Math Teacher
Helen	Bosch	Biology and Chemistry Teacher
Karla	Cramer	7th Grade Science/Math Teacher
Pam	Diamond	Language Arts Teacher
Maureen	Driscoll	Chemistry Instructor
Christine Renee	Floyd	Teacher
Beckie	Frisbee	Math teacher
Kris	Goyins	Teacher/Curriculum Spec.
Keith	Grebetz	English/Reading Teacher
Carmen	Hauck	Ibes and Chemistry Teacher
Jim	Hennig	Teacher
Estelle	Hill	High School Math Teacher
Rolland	Karlin	5th Grade Teacher
Nita	Kattell	5th Grade Teacher
Callie	Kolste	Supervising Teacher/K-2
Sarah	Norton	K-5 Special Education
Christine	Perkins	7th and 8th Grade Math Teacher
Connie	Sandvik	Middle School Counselor
Cherie	Stobie	Math Coach
Becky	Telling	Reading Instructor
Sandra	Wardell	Biology Teacher
Debra	Westrom	6th Grade Teacher

Appendix D—ITEM-LEVEL CLASSICAL STATISTICS

Table D-1. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 3

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
60919	MC	0.94	0.23	0
76772	MC	0.87	0.34	0
76759	MC	0.86	0.41	0
76864	MC	0.55	0.33	0
76840	MC	0.57	0.29	1
60313	MC	0.85	0.37	1
76781	MC	0.67	0.46	1
76795	MC	0.85	0.30	1
76750	MC	0.31	0.27	2
76895	MC	0.65	0.29	2
76911	MC	0.55	0.48	1
76988	MC	0.77	0.45	1
76751	MC	0.89	0.34	2
76884	MC	0.56	0.33	1
76904	MC	0.84	0.26	1
76782	MC	0.61	0.45	1
76979	MC	0.66	0.33	1
76855	MC	0.75	0.48	1
42994	SA	0.70	0.38	0
59293	SA	0.80	0.38	0
43261	CR	0.69	0.48	0
76756	MC	0.96	0.20	0
60918	MC	0.91	0.26	0
76917	MC	0.88	0.36	1
60974	MC	0.80	0.40	0
76879	MC	0.51	0.42	0
76765	MC	0.91	0.36	1
76769	MC	0.56	0.43	1
76752	MC	0.86	0.36	1
76913	MC	0.38	0.39	1
77003	MC	0.91	0.31	1
61065	MC	0.64	0.40	1
76859	MC	0.80	0.41	1
60952	MC	0.82	0.48	1
76774	MC	0.37	0.32	2
76909	MC	0.73	0.29	1
76841	MC	0.37	0.43	1
42983	MC	0.61	0.49	2
61046	MC	0.66	0.50	0
76866	MC	0.80	0.31	1
59294	SA	0.78	0.44	1
34561	MC	0.94	0.29	0
76842	MC	0.88	0.31	1
60278	MC	0.81	0.44	1
76906	MC	0.67	0.36	1
76784	MC	0.45	0.29	1
76915	MC	0.61	0.26	1
77019	MC	0.77	0.48	1
76971	MC	0.83	0.27	2

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
76843	MC	0.47	0.33	2
43090	MC	0.84	0.35	0
77006	MC	0.52	0.33	1
76777	MC	0.64	0.50	2
43136	MC	0.74	0.41	0
43105	MC	0.68	0.27	1
77011	MC	0.68	0.53	1
77027	MC	0.41	0.23	1
76836	MC	0.67	0.32	1
42962	MC	0.71	0.46	2
76899	CR	0.38	0.50	1

Table D-2. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 4

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
76972	MC	0.94	0.20	0
34588	MC	0.68	0.42	0
62320	MC	0.88	0.36	0
76823	MC	0.62	0.47	0
76812	MC	0.61	0.42	0
76961	MC	0.71	0.42	0
76892	MC	0.57	0.52	0
76939	MC	0.56	0.18	0
43173	MC	0.74	0.42	0
76788	MC	0.42	0.37	0
76941	MC	0.75	0.29	0
34965	MC	0.70	0.32	0
43312	MC	0.85	0.33	0
43296	MC	0.35	0.28	0
62363	MC	0.78	0.36	0
34594	MC	0.68	0.41	0
35215	MC	0.48	0.44	0
76969	MC	0.82	0.37	1
34900	SA	0.70	0.35	0
61780	SA	0.56	0.47	0
77063	CR	0.53	0.64	1
34952	MC	0.90	0.27	0
76959	MC	0.82	0.33	0
76888	MC	0.64	0.44	0
43332	MC	0.52	0.24	0
76834	MC	0.37	0.48	0
43314	MC	0.76	0.35	0
62214	MC	0.84	0.38	0
76844	MC	0.44	0.29	0
77022	MC	0.69	0.41	0
76794	MC	0.65	0.35	0
76935	MC	0.79	0.28	0
34602	MC	0.64	0.39	0
43330	MC	0.94	0.24	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
62401	MC	0.79	0.35	0
76924	MC	0.53	0.24	0
35792	MC	0.54	0.31	0
76830	MC	0.84	0.40	0
61804	MC	0.45	0.45	0
61829	MC	0.71	0.50	2
76763	SA	0.54	0.50	1
76948	MC	0.77	0.29	0
76819	MC	0.69	0.48	0
77050	MC	0.64	0.44	0
76827	MC	0.56	0.51	0
43298	MC	0.54	0.41	0
76995	MC	0.81	0.24	0
76856	MC	0.67	0.39	0
43369	MC	0.66	0.50	0
62389	MC	0.70	0.36	0
76965	MC	0.58	0.48	0
76933	MC	0.56	0.27	0
34744	MC	0.79	0.40	0
61811	MC	0.28	0.36	0
43167	MC	0.68	0.31	0
43276	MC	0.80	0.42	0
62381	MC	0.80	0.45	0
76963	MC	0.42	0.31	1
43320	MC	0.82	0.31	2
62486	CR	0.52	0.60	1

Table D-3. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 5

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
60551	MC	0.87	0.29	0
77210	MC	0.68	0.42	0
77270	MC	0.69	0.26	0
77314	MC	0.54	0.53	0
60417	MC	0.68	0.33	1
77325	MC	0.80	0.40	0
77247	MC	0.71	0.37	0
77259	MC	0.88	0.35	0
77179	MC	0.68	0.35	0
77274	MC	0.79	0.36	0
77193	MC	0.81	0.39	0
43514	MC	0.43	0.42	0
77228	MC	0.52	0.37	0
43560	MC	0.58	0.39	0
77245	MC	0.60	0.40	0
77205	MC	0.78	0.46	0
77310	MC	0.53	0.42	0
77220	MC	0.63	0.30	1
77298	SA	0.41	0.52	1

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
62025	SA	0.62	0.41	1
61052	CR	0.39	0.62	1
77200	MC	0.80	0.35	0
77163	MC	0.72	0.44	0
43477	MC	0.67	0.49	0
77318	MC	0.46	0.50	0
60370	MC	0.66	0.33	0
77257	MC	0.68	0.31	0
77204	MC	0.64	0.53	0
43471	MC	0.70	0.30	0
77208	MC	0.40	0.42	0
77181	MC	0.62	0.27	0
77191	MC	0.69	0.22	0
59800	MC	0.65	0.47	0
77243	MC	0.59	0.48	0
43409	MC	0.32	0.20	0
77279	MC	0.76	0.38	0
77207	MC	0.59	0.42	0
59986	MC	0.78	0.49	0
77330	MC	0.30	0.24	0
43413	MC	0.59	0.21	1
77295	SA	0.87	0.33	0
43516	MC	0.87	0.39	0
77219	MC	0.35	0.29	0
77321	MC	0.33	0.47	0
59856	MC	0.58	0.37	0
77186	MC	0.66	0.27	0
60504	MC	0.79	0.40	0
77265	MC	0.47	0.31	0
60371	MC	0.81	0.33	0
60971	MC	0.60	0.37	0
77217	MC	0.71	0.50	0
77198	MC	0.46	0.54	0
77255	MC	0.62	0.44	0
34658	MC	0.47	0.29	0
77177	MC	0.79	0.30	0
77282	MC	0.80	0.44	0
43585	MC	0.65	0.36	0
60911	MC	0.69	0.46	0
43421	MC	0.74	0.33	1
34660	CR	0.64	0.44	0

Table D-4. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 6

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
77377	MC	0.93	0.27	0
77320	MC	0.81	0.35	0
77376	MC	0.63	0.48	0
77340	MC	0.62	0.59	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
77323	MC	0.77	0.36	0
60885	MC	0.59	0.56	0
60901	MC	0.52	0.34	0
77378	MC	0.51	0.53	0
77317	MC	0.47	0.23	0
43912	MC	0.51	0.49	0
77373	MC	0.57	0.49	0
43863	MC	0.46	0.32	0
43921	MC	0.56	0.32	0
77313	MC	0.42	0.35	1
63005	SA	0.41	0.41	0
44088	SA	0.86	0.34	1
77642	SA	0.49	0.57	1
77649	CR	0.25	0.63	5
44004	MC	0.77	0.30	0
62994	MC	0.45	0.35	0
77522	MC	0.64	0.39	0
34842	MC	0.73	0.39	0
77630	MC	0.57	0.49	0
43992	MC	0.52	0.19	0
77614	MC	0.75	0.31	0
77339	MC	0.51	0.20	0
77476	MC	0.81	0.20	0
77553	MC	0.53	0.41	0
77451	MC	0.49	0.38	0
77398	MC	0.40	0.26	0
77608	MC	0.46	0.32	0
61156	MC	0.47	0.50	0
77515	MC	0.50	0.38	0
77509	MC	0.69	0.30	0
77538	MC	0.69	0.42	0
62050	MC	0.38	0.31	0
77414	MC	0.73	0.31	0
77542	MC	0.35	0.42	0
43995	MC	0.80	0.22	0
77445	MC	0.95	0.17	0
77621	MC	0.76	0.47	0
43956	MC	0.63	0.39	0
77502	MC	0.63	0.40	0
77582	MC	0.66	0.36	0
62060	MC	0.85	0.42	0
77380	MC	0.54	0.47	0
43944	MC	0.67	0.44	0
34546	MC	0.47	0.43	0
77449	MC	0.55	0.35	0
43949	MC	0.48	0.30	0
62029	MC	0.64	0.32	0
77625	MC	0.87	0.39	0
77633	MC	0.60	0.34	0
43991	MC	0.50	0.28	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
43963	MC	0.70	0.47	0
34913	MC	0.47	0.46	0
77455	MC	0.70	0.42	0
77497	MC	0.45	0.19	0
77517	MC	0.62	0.42	1
44048	CR	0.41	0.61	1

Table D-5. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 7

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
61206	MC	0.78	0.35	0
86297	MC	0.74	0.49	0
61204	MC	0.49	0.37	0
86366	MC	0.56	0.41	0
86305	MC	0.40	0.32	0
61205	MC	0.38	0.37	0
43836	MC	0.60	0.38	0
62948	MC	0.37	0.46	0
86280	MC	0.78	0.44	0
86296	MC	0.56	0.44	0
86300	MC	0.33	0.31	0
86302	MC	0.51	0.43	0
61228	MC	0.42	0.43	1
86295	MC	0.67	0.46	1
86348	SA	0.63	0.56	1
43799	SA	0.73	0.47	1
86349	SA	0.55	0.59	1
158633	CR	0.39	0.61	7
61742	MC	0.54	0.22	0
86431	MC	0.42	0.29	0
43787	MC	0.62	0.36	0
86438	MC	0.95	0.16	0
86382	MC	0.79	0.40	0
86615	MC	0.42	0.45	0
86555	MC	0.75	0.38	0
43721	MC	0.56	0.33	0
86549	MC	0.52	0.45	0
43782	MC	0.49	0.33	0
86455	MC	0.60	0.46	0
86635	MC	0.87	0.35	0
86395	MC	0.45	0.47	0
61745	MC	0.66	0.47	0
86374	MC	0.60	0.43	0
61279	MC	0.87	0.43	0
86631	MC	0.50	0.49	0
61195	MC	0.36	0.32	0
43705	MC	0.23	0.19	0
86381	MC	0.45	0.39	0
86379	MC	0.66	0.42	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
86535	MC	0.79	0.27	0
86369	MC	0.68	0.48	0
86448	MC	0.53	0.21	0
86473	MC	0.66	0.39	0
86692	MC	0.62	0.42	0
61766	MC	0.88	0.36	0
86458	MC	0.55	0.48	0
86650	MC	0.77	0.42	0
61250	MC	0.72	0.40	0
86683	MC	0.74	0.41	0
86622	MC	0.57	0.43	0
86311	MC	0.46	0.42	0
86313	MC	0.59	0.39	0
86453	MC	0.37	0.46	0
86591	MC	0.84	0.35	0
86482	MC	0.72	0.38	0
86568	MC	0.49	0.33	0
86689	MC	0.63	0.56	0
86675	MC	0.44	0.36	0
43700	MC	0.51	0.32	1
43922	CR	0.43	0.67	1

Table D-6. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 8

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
87598	MC	0.97	0.20	0
63025	MC	0.45	0.35	0
87593	MC	0.77	0.37	0
87658	MC	0.53	0.44	0
87527	MC	0.65	0.38	0
87583	MC	0.68	0.22	0
87808	MC	0.38	0.43	0
87802	MC	0.64	0.48	0
87661	MC	0.31	0.21	0
87580	MC	0.66	0.20	0
44214	MC	0.50	0.30	0
87606	MC	0.55	0.37	0
44176	MC	0.34	0.25	1
44199	SA	0.52	0.51	1
87834	SA	0.57	0.46	1
87841	SA	0.50	0.52	1
44267	CR	0.41	0.58	4
44626	MC	0.90	0.26	0
88363	MC	0.43	0.30	0
88177	MC	0.76	0.33	0
44160	MC	0.46	0.25	0
88019	MC	0.63	0.21	0
88864	MC	0.64	0.30	0
87623	MC	0.31	0.28	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
44137	MC	0.60	0.33	0
44141	MC	0.80	0.40	0
88838	MC	0.64	0.31	0
63148	MC	0.27	0.25	0
63219	MC	0.56	0.34	0
88174	MC	0.55	0.23	0
88189	MC	0.31	0.27	0
63144	MC	0.44	0.30	0
62943	MC	0.57	0.32	0
44642	MC	0.18	0.23	0
88338	MC	0.56	0.38	0
44662	MC	0.48	0.43	0
44220	MC	0.55	0.46	0
88216	MC	0.72	0.38	0
63038	MC	0.82	0.45	0
87799	MC	0.54	0.38	0
34945	MC	0.43	0.37	0
35019	MC	0.88	0.39	0
44205	MC	0.75	0.42	0
63252	MC	0.46	0.26	0
63203	MC	0.43	0.37	0
44123	MC	0.50	0.23	0
63115	MC	0.46	0.26	0
44210	MC	0.65	0.37	0
44648	MC	0.57	0.47	0
88086	MC	0.52	0.36	0
88848	MC	0.61	0.40	0
63287	MC	0.67	0.35	0
88263	MC	0.59	0.42	0
63106	MC	0.85	0.30	0
63256	MC	0.72	0.36	0
86642	MC	0.31	0.33	0
88381	MC	0.63	0.38	0
88352	MC	0.69	0.34	0

Table D-7. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Mathematics Grade 10

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
59397	MC	0.72	0.39	0
77480	MC	0.51	0.28	1
77571	MC	0.40	0.39	0
61319	MC	0.44	0.50	0
77612	MC	0.60	0.38	0
61298	MC	0.31	0.25	0
77623	MC	0.59	0.46	0
77618	MC	0.76	0.40	0
77570	MC	0.45	0.49	0
59377	MC	0.53	0.40	0
77596	MC	0.56	0.30	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
43613	MC	0.31	0.25	0
59365	MC	0.30	0.46	0
34804	MC	0.51	0.30	0
43899	SA	0.36	0.44	1
34864	SA	0.35	0.33	8
77635	SA	0.54	0.43	1
160523	CR	0.33	0.57	9
77368	MC	0.59	0.33	0
43969	MC	0.88	0.29	0
62368	MC	0.61	0.36	0
77507	MC	0.44	0.37	0
77352	MC	0.22	0.26	0
77371	MC	0.79	0.38	0
77485	MC	0.54	0.50	0
34881	MC	0.46	0.30	0
77392	MC	0.56	0.47	0
43743	MC	0.27	0.30	0
34835	MC	0.38	0.27	0
77382	MC	0.35	0.36	0
43822	MC	0.66	0.37	0
43917	MC	0.51	0.39	0
34685	MC	0.74	0.45	0
62292	MC	0.39	0.21	0
77384	MC	0.31	0.25	0
77432	MC	0.68	0.23	0
77503	MC	0.43	0.12	0
62202	MC	0.56	0.33	0
34838	MC	0.62	0.43	0
62177	MC	0.75	0.24	0
77484	MC	0.69	0.27	1
77370	MC	0.58	0.25	0
77354	MC	0.60	0.19	0
77415	MC	0.53	0.32	0
77562	MC	0.31	0.35	0
61265	MC	0.37	0.36	0
77428	MC	0.65	0.41	0
77561	MC	0.54	0.41	0
62333	MC	0.63	0.35	0
77394	MC	0.36	0.38	0
61324	MC	0.88	0.34	0
35234	MC	0.25	0.14	0
77551	MC	0.53	0.30	0
77619	MC	0.45	0.45	0
61281	MC	0.70	0.45	0
62286	MC	0.31	0.28	0
34856	MC	0.44	0.38	0
34471	MC	0.31	0.37	1
61312	MC	0.63	0.45	0
159645	CR	0.25	0.56	3

Table D-8. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 3

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
68808	MC	0.85	0.37	0
68809	MC	0.77	0.41	0
68811	MC	0.95	0.35	0
68810	MC	0.54	0.36	1
68812	MC	0.53	0.37	0
68814	MC	0.48	0.22	1
68818	MC	0.66	0.26	3
92739	MC	0.81	0.46	0
92742	MC	0.69	0.41	0
92743	MC	0.69	0.48	1
92745	MC	0.54	0.35	0
92746	MC	0.89	0.43	1
92748	MC	0.87	0.33	1
92749	MC	0.79	0.45	0
92750	MC	0.75	0.42	0
92751	MC	0.75	0.30	1
92758	MC	0.80	0.37	1
92752	MC	0.61	0.28	1
92755	MC	0.82	0.42	1
92761	CR	0.33	0.50	1
92786	MC	0.61	0.28	0
92789	MC	0.92	0.37	0
92791	MC	0.84	0.28	0
92792	MC	0.81	0.31	0
92794	MC	0.77	0.43	1
92797	MC	0.76	0.49	0
92795	MC	0.60	0.38	1
92670	MC	0.91	0.38	0
92673	MC	0.59	0.36	0
92674	MC	0.57	0.35	1
92675	MC	0.85	0.46	1
92677	MC	0.63	0.33	1
92695	MC	0.80	0.43	1
92696	MC	0.56	0.38	2
92798	MC	0.85	0.35	0
92800	MC	0.75	0.56	0
157468	MC	0.61	0.34	0
92802	MC	0.51	0.48	1
92803	MC	0.77	0.42	0
92808	MC	0.69	0.37	1
157469	MC	0.51	0.25	2
92763	MC	0.73	0.45	0
92765	MC	0.54	0.36	1
92767	MC	0.82	0.46	1
92771	MC	0.61	0.35	1
92768	MC	0.84	0.40	1
92766	MC	0.69	0.41	1
92773	MC	0.72	0.41	0
92775	MC	0.88	0.40	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
92777	MC	0.57	0.41	1
92781	MC	0.57	0.36	1
92779	MC	0.57	0.32	1
92778	MC	0.68	0.47	3
92783	CR	0.29	0.49	1

Table D-9. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 4

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
93939	MC	0.69	0.37	0
157543	MC	0.84	0.34	0
93944	MC	0.80	0.47	0
93946	MC	0.80	0.33	0
93948	MC	0.69	0.26	0
93951	MC	0.83	0.29	0
157544	MC	0.60	0.33	0
67330	MC	0.77	0.32	0
67333	MC	0.63	0.43	0
67334	MC	0.57	0.45	0
67365	MC	0.73	0.49	0
67350	MC	0.76	0.42	0
67354	MC	0.52	0.38	0
67346	MC	0.75	0.33	0
67359	MC	0.59	0.41	0
67367	MC	0.67	0.41	0
67371	MC	0.74	0.45	1
67374	MC	0.68	0.38	0
67368	MC	0.69	0.43	1
67382	CR	0.35	0.45	1
93837	MC	0.82	0.39	0
93838	MC	0.71	0.34	0
93846	MC	0.67	0.44	0
93876	MC	0.84	0.39	0
93881	MC	0.67	0.49	0
93895	MC	0.87	0.44	0
93888	MC	0.55	0.35	0
93842	MC	0.56	0.30	0
93855	MC	0.87	0.36	0
93857	MC	0.67	0.37	0
93862	MC	0.75	0.40	0
93863	MC	0.76	0.46	0
93849	MC	0.69	0.31	0
93874	MC	0.68	0.35	0
94002	MC	0.81	0.37	0
94004	MC	0.77	0.40	0
157546	MC	0.71	0.41	0
94016	MC	0.64	0.35	0
94015	MC	0.81	0.48	0
94027	MC	0.69	0.47	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
157548	MC	0.51	0.09	0
94048	MC	0.74	0.45	0
94050	MC	0.72	0.43	0
94072	MC	0.75	0.52	0
94083	MC	0.58	0.34	0
94079	MC	0.88	0.46	0
94108	MC	0.78	0.37	0
94092	MC	0.86	0.41	0
94095	MC	0.69	0.46	0
94077	MC	0.77	0.50	0
94111	MC	0.76	0.39	0
94120	MC	0.75	0.47	0
94116	MC	0.46	0.34	0
94139	CR	0.37	0.46	1

Table D-10. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 5

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
65312	MC	0.77	0.47	0
65317	MC	0.73	0.31	0
65383	MC	0.82	0.30	0
65368	MC	0.70	0.37	0
69235	MC	0.80	0.36	0
65379	MC	0.82	0.28	0
65387	MC	0.41	0.36	0
93601	MC	0.79	0.37	0
93611	MC	0.67	0.19	0
93612	MC	0.63	0.33	0
93615	MC	0.85	0.49	0
93616	MC	0.72	0.45	0
93626	MC	0.61	0.36	0
93631	MC	0.82	0.42	0
93623	MC	0.75	0.46	0
93628	MC	0.68	0.29	0
93635	MC	0.69	0.24	0
93639	MC	0.90	0.43	0
93638	MC	0.69	0.34	0
93668	CR	0.47	0.51	0
93353	MC	0.59	0.35	0
93366	MC	0.73	0.29	0
93375	MC	0.90	0.32	0
93378	MC	0.76	0.43	0
93381	MC	0.86	0.38	0
93389	MC	0.64	0.36	0
93385	MC	0.59	0.33	0
93529	MC	0.79	0.42	0
93510	MC	0.69	0.48	0
93533	MC	0.58	0.35	0
93524	MC	0.71	0.46	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
93520	MC	0.68	0.40	0
93536	MC	0.76	0.40	0
93537	MC	0.63	0.41	0
93700	MC	0.90	0.33	0
93695	MC	0.69	0.36	0
93698	MC	0.82	0.42	0
93705	MC	0.54	0.40	0
93704	MC	0.79	0.40	0
93709	MC	0.62	0.40	0
93711	MC	0.83	0.40	0
93448	MC	0.78	0.38	0
93414	MC	0.82	0.49	0
93415	MC	0.79	0.38	0
93416	MC	0.76	0.50	0
93419	MC	0.90	0.39	0
93420	MC	0.81	0.36	0
93431	MC	0.64	0.36	0
93421	MC	0.81	0.47	0
93428	MC	0.84	0.54	0
93446	MC	0.49	0.28	0
93444	MC	0.75	0.40	0
93451	MC	0.49	0.34	0
93462	CR	0.49	0.40	0

Table D-11. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 6

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
95410	MC	0.74	0.31	0
95421	MC	0.85	0.44	0
95409	MC	0.80	0.32	0
95445	MC	0.58	0.37	0
95431	MC	0.65	0.40	0
95435	MC	0.65	0.31	0
95450	MC	0.77	0.29	0
95305	MC	0.63	0.28	0
95335	MC	0.68	0.25	0
95330	MC	0.72	0.42	0
95351	MC	0.76	0.43	0
95358	MC	0.84	0.44	0
95363	MC	0.88	0.44	0
95369	MC	0.83	0.35	0
95371	MC	0.84	0.49	0
95375	MC	0.76	0.48	0
95386	MC	0.53	0.36	1
95381	MC	0.61	0.38	0
95393	MC	0.64	0.37	0
95397	CR	0.49	0.48	0
95077	MC	0.55	0.32	0
95088	MC	0.89	0.36	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
95092	MC	0.84	0.32	0
95101	MC	0.86	0.35	0
95132	MC	0.64	0.39	0
95105	MC	0.61	0.33	0
95115	MC	0.72	0.41	0
95348	MC	0.68	0.32	0
95342	MC	0.77	0.31	0
95353	MC	0.62	0.41	0
95368	MC	0.90	0.39	0
97773	MC	0.73	0.38	0
95387	MC	0.66	0.33	0
95398	MC	0.82	0.39	0
95202	MC	0.63	0.45	0
95183	MC	0.41	0.29	0
95218	MC	0.70	0.50	0
95228	MC	0.80	0.42	0
95231	MC	0.75	0.47	0
95289	MC	0.53	0.23	0
95299	MC	0.49	0.30	0
95033	MC	0.81	0.34	0
95036	MC	0.85	0.32	0
95041	MC	0.53	0.34	0
95045	MC	0.79	0.46	0
95085	MC	0.91	0.36	0
95089	MC	0.62	0.42	0
95093	MC	0.80	0.42	0
95121	MC	0.60	0.41	0
95110	MC	0.77	0.42	0
95114	MC	0.76	0.44	0
95145	MC	0.66	0.41	0
95157	MC	0.83	0.48	0
95171	CR	0.53	0.46	0

Table D-12. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 7

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
41859	MC	0.60	0.34	0
157578	MC	0.57	0.39	0
41860	MC	0.74	0.37	0
41867	MC	0.80	0.29	0
41864	MC	0.85	0.49	0
41866	MC	0.76	0.40	0
41868	MC	0.85	0.47	0
92567	MC	0.81	0.47	0
157579	MC	0.85	0.52	0
92583	MC	0.67	0.33	0
92588	MC	0.73	0.35	0
92589	MC	0.71	0.50	0
92591	MC	0.88	0.46	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
92593	MC	0.76	0.38	0
92596	MC	0.58	0.31	0
92600	MC	0.82	0.45	0
92605	MC	0.79	0.40	0
92606	MC	0.59	0.45	0
92608	MC	0.82	0.45	0
92611	CR	0.52	0.52	0
68610	MC	0.90	0.36	0
68611	MC	0.66	0.31	0
68612	MC	0.79	0.32	0
68616	MC	0.79	0.40	0
68613	MC	0.63	0.44	0
68614	MC	0.56	0.41	0
68620	MC	0.88	0.43	0
92341	MC	0.63	0.37	0
92342	MC	0.64	0.49	0
92343	MC	0.75	0.38	0
92345	MC	0.76	0.43	0
92348	MC	0.69	0.29	0
92350	MC	0.75	0.40	0
92347	MC	0.65	0.34	0
68493	MC	0.83	0.40	0
68495	MC	0.60	0.41	0
68497	MC	0.75	0.49	0
68510	MC	0.83	0.49	0
68507	MC	0.88	0.40	0
68498	MC	0.81	0.34	0
68514	MC	0.73	0.32	0
92531	MC	0.61	0.33	0
92535	MC	0.76	0.51	0
92536	MC	0.64	0.41	0
92540	MC	0.56	0.29	0
92541	MC	0.63	0.50	0
92543	MC	0.70	0.39	0
92545	MC	0.62	0.44	0
92549	MC	0.84	0.53	0
92554	MC	0.81	0.51	0
92555	MC	0.67	0.43	0
92558	MC	0.53	0.28	0
92559	MC	0.55	0.40	0
92562	CR	0.47	0.58	0

Table D-13. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 8

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
67937	MC	0.70	0.32	0
67938	MC	0.78	0.32	0
67944	MC	0.81	0.35	0
67948	MC	0.63	0.28	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
67953	MC	0.54	0.30	0
67952	MC	0.88	0.28	0
67966	MC	0.83	0.30	0
95838	MC	0.74	0.35	0
95843	MC	0.75	0.35	0
95844	MC	0.62	0.41	0
95845	MC	0.78	0.34	0
95847	MC	0.73	0.31	0
95851	MC	0.84	0.32	0
95853	MC	0.78	0.30	0
95855	MC	0.70	0.28	0
95856	MC	0.84	0.41	0
95863	MC	0.87	0.45	0
95867	MC	0.84	0.44	0
95866	MC	0.65	0.25	0
95869	CR	0.57	0.52	0
68698	MC	0.69	0.39	0
68699	MC	0.77	0.47	0
68702	MC	0.67	0.43	0
68714	MC	0.72	0.35	0
68725	MC	0.80	0.40	0
68724	MC	0.79	0.32	0
68726	MC	0.85	0.47	0
95688	MC	0.73	0.45	0
95691	MC	0.77	0.45	0
95700	MC	0.67	0.49	0
95703	MC	0.71	0.46	0
157551	MC	0.60	0.50	0
95706	MC	0.83	0.23	0
95708	MC	0.70	0.42	0
95604	MC	0.79	0.28	0
95627	MC	0.83	0.46	0
95637	MC	0.80	0.54	0
95644	MC	0.75	0.46	0
95647	MC	0.69	0.43	0
95649	MC	0.66	0.38	0
95656	MC	0.78	0.41	0
68470	MC	0.69	0.45	0
68473	MC	0.74	0.36	0
68476	MC	0.77	0.44	0
68478	MC	0.82	0.43	0
68475	MC	0.78	0.51	0
68480	MC	0.68	0.39	0
68487	MC	0.66	0.31	0
68500	MC	0.80	0.54	0
68499	MC	0.77	0.37	0
68504	MC	0.75	0.47	0
68508	MC	0.63	0.38	0
68501	MC	0.73	0.44	0
68511	CR	0.51	0.44	1

Table D-14. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Reading Grade 10

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
67599	MC	0.87	0.32	0
67687	MC	0.83	0.49	0
67720	MC	0.54	0.37	0
67741	MC	0.62	0.39	0
67750	MC	0.79	0.35	0
67753	MC	0.78	0.43	0
67757	MC	0.74	0.35	0
94842	MC	0.65	0.28	0
94843	MC	0.87	0.38	0
94848	MC	0.78	0.30	0
94852	MC	0.87	0.43	0
94854	MC	0.87	0.44	0
94857	MC	0.91	0.49	0
94861	MC	0.76	0.47	0
94874	MC	0.73	0.47	0
94877	MC	0.78	0.41	0
94879	MC	0.69	0.40	0
94863	MC	0.70	0.27	0
94882	MC	0.55	0.39	0
94887	CR	0.54	0.53	1
95338	MC	0.77	0.48	0
95340	MC	0.76	0.40	0
95361	MC	0.76	0.42	0
95367	MC	0.71	0.44	0
95374	MC	0.65	0.36	0
95377	MC	0.88	0.43	0
95391	MC	0.67	0.28	0
95026	MC	0.81	0.34	0
95030	MC	0.86	0.41	0
95138	MC	0.72	0.48	0
95164	MC	0.78	0.34	0
95187	MC	0.61	0.22	0
95154	MC	0.86	0.43	0
95207	MC	0.93	0.41	0
95216	MC	0.75	0.46	0
95273	MC	0.63	0.35	0
95234	MC	0.74	0.42	0
95279	MC	0.64	0.38	0
95285	MC	0.67	0.30	0
95290	MC	0.59	0.33	0
95293	MC	0.75	0.43	0
94889	MC	0.83	0.44	0
94890	MC	0.63	0.45	0
94892	MC	0.71	0.28	0
94894	MC	0.62	0.42	0
94898	MC	0.65	0.34	0
94903	MC	0.72	0.38	0
94912	MC	0.71	0.40	0
94924	MC	0.64	0.41	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
94929	MC	0.79	0.45	0
94931	MC	0.81	0.43	1
94941	MC	0.83	0.43	0
94943	MC	0.81	0.41	0
94955	CR	0.46	0.57	3

Table D-15. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Science Grade 4

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
39067	MC	0.55	0.29	0
75824	MC	0.75	0.25	0
75690	MC	0.57	0.30	0
39242	MC	0.62	0.17	0
39314	MC	0.53	0.30	0
57874	MC	0.87	0.29	0
53659	MC	0.88	0.29	0
75923	MC	0.51	0.31	0
75835	MC	0.65	0.29	0
39229	MC	0.60	0.34	0
39086	MC	0.60	0.32	0
76403	MC	0.83	0.30	0
39336	MC	0.37	0.31	0
42786	MC	0.50	0.19	0
76285	MC	0.82	0.40	0
39184	MC	0.70	0.37	0
75889	MC	0.51	0.41	1
42790	CR	0.65	0.46	1
75902	MC	0.72	0.21	0
39119	MC	0.64	0.26	0
76394	MC	0.76	0.27	0
75784	MC	0.53	0.39	0
75910	MC	0.87	0.26	0
39127	MC	0.52	0.21	0
39257	MC	0.67	0.33	0
39238	MC	0.63	0.30	0
76296	MC	0.77	0.29	0
75717	MC	0.69	0.25	0
75418	MC	0.69	0.20	0
75737	MC	0.79	0.44	0
75788	MC	0.89	0.28	0
75833	MC	0.85	0.37	0
75741	MC	0.66	0.36	0
75912	MC	0.71	0.35	0
75908	MC	0.73	0.17	0
75702	MC	0.89	0.23	1
55629	MC	0.36	0.25	0
75421	MC	0.51	0.25	0
75887	MC	0.79	0.37	0
75694	MC	0.64	0.26	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
53393	MC	0.61	0.23	0
75828	MC	0.72	0.45	0
75801	MC	0.57	0.41	0
75782	MC	0.46	0.19	0
75901	MC	0.67	0.38	0
75895	MC	0.57	0.38	0
75899	MC	0.72	0.44	0
75408	MC	0.54	0.20	0
55464	MC	0.87	0.24	0
57860	MC	0.63	0.41	0
75517	MC	0.59	0.31	0
56340	MC	0.61	0.32	0
75423	MC	0.84	0.28	0
75752	MC	0.48	0.21	1
75427	CR	0.33	0.44	1

Table D-16. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Science Grade 8

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
89277	MC	0.93	0.35	0
89520	MC	0.84	0.28	0
39701	MC	0.71	0.26	0
89817	MC	0.69	0.21	0
89693	MC	0.61	0.34	0
39782	MC	0.55	0.48	0
39745	MC	0.74	0.44	0
89647	MC	0.75	0.29	0
89361	MC	0.67	0.35	0
89911	MC	0.64	0.19	0
89593	MC	0.74	0.33	0
56851	MC	0.43	0.27	0
89582	MC	0.35	0.20	0
89850	MC	0.48	0.15	0
89522	MC	0.65	0.35	0
89892	MC	0.68	0.40	0
89639	MC	0.47	0.34	0
55106	CR	0.71	0.50	1
56805	MC	0.84	0.30	0
89420	MC	0.82	0.31	0
89263	MC	0.73	0.42	0
39659	MC	0.47	0.31	0
56828	MC	0.44	0.34	0
89498	MC	0.66	0.31	0
89778	MC	0.48	0.23	0
89652	MC	0.40	0.28	0
89795	MC	0.71	0.44	0
89742	MC	0.38	0.22	0
89691	MC	0.57	0.39	0
89457	MC	0.45	0.18	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
56897	MC	0.68	0.41	0
89884	MC	0.62	0.45	0
89468	MC	0.48	0.31	0
89634	MC	0.47	0.38	0
89452	MC	0.70	0.25	0
89766	MC	0.61	0.32	0
38602	MC	0.82	0.38	0
89752	MC	0.73	0.32	0
54264	MC	0.78	0.38	0
89770	MC	0.54	0.37	0
89444	MC	0.38	0.25	0
89726	MC	0.45	0.35	0
39652	MC	0.48	0.44	0
56833	MC	0.46	0.30	0
89849	MC	0.65	0.43	0
89508	MC	0.48	0.30	0
56992	MC	0.37	0.27	0
54543	MC	0.60	0.44	0
39780	MC	0.85	0.38	0
89505	MC	0.38	0.35	0
54454	MC	0.63	0.31	0
89870	MC	0.75	0.45	0
89382	MC	0.72	0.32	0
89863	MC	0.51	0.37	0
39764	CR	0.52	0.53	1

Table D-17. 2009–10 MontCAS: Item Level Classical Test Theory Statistics – Science Grade 10

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
40317	MC	0.78	0.41	0
55819	MC	0.75	0.32	0
75948	MC	0.49	0.28	0
75876	MC	0.60	0.30	0
60856	MC	0.63	0.33	0
53812	MC	0.50	0.36	0
75445	MC	0.38	0.22	0
75650	MC	0.73	0.31	0
75958	MC	0.42	0.23	0
75433	MC	0.52	0.42	0
53584	MC	0.41	0.23	0
55620	MC	0.65	0.43	0
75440	MC	0.49	0.40	0
75780	MC	0.65	0.31	0
53265	MC	0.72	0.43	0
75456	MC	0.56	0.27	0
75634	MC	0.44	0.46	0
75461	CR	0.45	0.58	4
40331	MC	0.83	0.28	0
75739	MC	0.70	0.39	0

continued

Item number	*Item type*	*Difficulty*	*Discrimination*	*Percent omitted*
75963	MC	0.60	0.34	0
75611	MC	0.62	0.27	0
75880	MC	0.63	0.35	0
75629	MC	0.67	0.43	0
75859	MC	0.58	0.34	0
75764	MC	0.84	0.42	0
75635	MC	0.52	0.28	0
75807	MC	0.52	0.32	0
75442	MC	0.54	0.40	0
75856	MC	0.77	0.43	0
56702	MC	0.56	0.35	0
75941	MC	0.71	0.35	0
75785	MC	0.71	0.29	0
75861	MC	0.51	0.33	0
75878	MC	0.33	0.23	0
75706	MC	0.65	0.49	0
40285	MC	0.74	0.29	0
55696	MC	0.71	0.15	0
54221	MC	0.67	0.28	0
75873	MC	0.50	0.19	0
75980	MC	0.53	0.22	0
55710	MC	0.58	0.38	0
75620	MC	0.36	0.29	0
75972	MC	0.59	0.36	0
75804	MC	0.47	0.40	0
56698	MC	0.48	0.19	0
75631	MC	0.55	0.41	0
56234	MC	0.48	0.34	0
75863	MC	0.64	0.48	0
52985	MC	0.53	0.27	0
75854	MC	0.63	0.31	0
75970	MC	0.54	0.38	0
75638	MC	0.73	0.36	0
75937	MC	0.80	0.47	0
75652	CR	0.16	0.49	4

Appendix E—Item-Level Score Distributions

Table E-1. 2009–10 MontCAS: Item-Level Score Distributions for Constructed-Response Items by Subject and Grade

Subject	*Grade*	*Item number*	*Total possible points*	*Percent of students at score point*				
				0	*1*	*2*	*3*	*4*
Mathematics	3	43261	4	3	6	21	51	18
		76899	4	20	28	30	19	2
	4	77063	4	8	29	14	34	13
		62486	4	30	12	9	14	34
	5	61052	4	27	26	14	23	8
		34660	4	10	11	28	16	35
	6	77649	4	52	10	19	6	9
		44048	4	20	32	19	20	8
	7	158633	4	24	20	20	22	8
		43922	4	29	16	28	7	20
	8	44267	4	23	25	23	11	16
	10	160523	4	33	11	27	11	9
		159645	4	36	29	26	5	1
Reading	3	92761	4	18	35	42	4	0
		92783	4	26	40	25	7	2
	4	67382	4	17	36	36	8	2
		94139	4	12	39	40	8	1
	5	93668	4	7	28	38	19	7
		93462	4	1	23	56	17	2
	6	95397	4	10	21	36	24	8
		95171	4	2	21	46	23	7
	7	92611	4	4	21	43	27	6
		92562	4	7	27	40	21	4
	8	95869	4	2	22	36	26	14
		68511	4	3	27	38	20	10
	10	94887	4	3	17	43	32	5
		94955	4	5	29	38	21	5
Science	4	42790	4	7	17	16	27	33
		75427	4	19	38	35	5	2
	8	55106	4	4	8	18	34	35
		39764	4	30	10	14	12	33
	10	75461	4	18	22	21	27	9
		75652	4	50	31	14	2	0

Appendix F—Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored

Table F-1. 2009–10 MontCAS: Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored – Mathematics

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
3	Male	Female	MC	55	2	1	1	0	0	0
	White	Hispanic	MC	55	6	5	1	0	0	0
	White	Native American	MC	55	5	5	0	0	0	0
	No Disability	Disability	MC	55	1	1	0	0	0	0
	Not Low Income	Low Income	MC	55	1	1	0	0	0	0
	Male	Female	OR	5	1	1	0	0	0	0
	White	Hispanic	OR	5	1	0	1	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	1	1	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
4	Male	Female	MC	55	5	4	1	0	0	0
	White	Hispanic	MC	55	2	1	1	0	0	0
	White	Native American	MC	55	7	7	0	0	0	0
	No Disability	Disability	MC	55	2	2	0	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Male	Female	OR	5	0	0	0	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	0	0	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
5	Male	Female	MC	55	8	8	0	0	0	0
	White	Hispanic	MC	55	5	4	1	0	0	0
	White	Native American	MC	55	3	3	0	0	0	0
	No Disability	Disability	MC	55	5	4	1	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Male	Female	OR	5	1	0	1	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	0	0	0	0	0	0
	No Disability	Disability	OR	5	1	1	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
6	Male	Female	MC	55	6	3	3	1	1	0
	White	Hispanic	MC	55	4	3	1	0	0	0
	White	Native American	MC	55	2	1	1	0	0	0
	No Disability	Disability	MC	55	7	7	0	0	0	0

continued

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
6	Not Low Income	Low Income	MC	55	1	1	0	0	0	0
	Male	Female	OR	5	3	1	2	0	0	0
	White	Hispanic	OR	5	1	1	0	0	0	0
	White	Native American	OR	5	1	1	0	0	0	0
	No Disability	Disability	OR	5	1	1	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
7	Male	Female	MC	55	6	5	1	0	0	0
	White	Hispanic	MC	55	7	4	3	0	0	0
	White	Native American	MC	55	1	1	0	0	0	0
	No Disability	Disability	MC	55	8	7	1	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Male	Female	OR	5	0	0	0	0	0	0
	White	Hispanic	OR	5	0	0	0	0	0	0
	White	Native American	OR	5	2	2	0	0	0	0
	No Disability	Disability	OR	5	2	2	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0
8	Male	Female	MC	54	8	4	4	0	0	0
	White	Hispanic	MC	54	4	4	0	0	0	0
	White	Native American	MC	54	4	4	0	0	0	0
	No Disability	Disability	MC	54	9	9	0	1	1	0
	Not Low Income	Low Income	MC	54	0	0	0	0	0	0
	Male	Female	OR	4	1	1	0	0	0	0
	White	Hispanic	OR	4	0	0	0	0	0	0
	White	Native American	OR	4	1	1	0	0	0	0
	No Disability	Disability	OR	4	0	0	0	0	0	0
	Not Low Income	Low Income	OR	4	0	0	0	0	0	0
10	Male	Female	MC	55	5	3	2	2	2	0
	White	Hispanic	MC	55	3	1	2	1	0	1
	White	Native American	MC	55	5	5	0	0	0	0
	No Disability	Disability	MC	55	14	12	2	0	0	0
	Not Low Income	Low Income	MC	55	0	0	0	0	0	0
	Male	Female	OR	5	0	0	0	0	0	0
	White	Hispanic	OR	5	1	1	0	0	0	0
	White	Native American	OR	5	1	0	1	0	0	0
	No Disability	Disability	OR	5	3	3	0	0	0	0
	Not Low Income	Low Income	OR	5	0	0	0	0	0	0

Table F-2. 2009–10 MontCAS: Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored – Reading

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
3	Male	Female	MC	52	4	1	3	0	0	0
	White	Hispanic	MC	52	2	0	2	0	0	0
	White	Native American	MC	52	1	0	1	1	0	1
	No Disability	Disability	MC	52	4	1	3	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Male	Female	OR	2	0	0	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
4	Male	Female	MC	52	3	0	3	0	0	0
	White	Hispanic	MC	52	2	0	2	0	0	0
	White	Native American	MC	52	2	0	2	0	0	0
	No Disability	Disability	MC	52	2	0	2	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Male	Female	OR	2	0	0	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
5	Male	Female	MC	52	8	3	5	0	0	0
	White	Hispanic	MC	52	5	2	3	0	0	0
	White	Native American	MC	52	3	0	3	1	0	1
	No Disability	Disability	MC	52	3	1	2	1	0	1
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Male	Female	OR	2	1	1	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
6	Male	Female	MC	52	7	1	6	1	0	1
	White	Hispanic	MC	52	3	0	3	1	0	1
	White	Native American	MC	52	7	0	7	0	0	0
	No Disability	Disability	MC	52	1	0	1	0	0	0

continued

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
6	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Male	Female	OR	2	2	2	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	2	0	2	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
7	Male	Female	MC	52	4	1	3	0	0	0
	White	Hispanic	MC	52	3	0	3	0	0	0
	White	Native American	MC	52	2	1	1	0	0	0
	No Disability	Disability	MC	52	7	0	7	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Male	Female	OR	2	2	2	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	2	0	2	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
8	Male	Female	MC	52	2	2	0	1	0	1
	White	Hispanic	MC	52	0	0	0	0	0	0
	White	Native American	MC	52	3	0	3	0	0	0
	No Disability	Disability	MC	52	3	1	2	0	0	0
	Not Low Income	Low Income	MC	52	0	0	0	0	0	0
	Male	Female	OR	2	2	2	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	2	0	2	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
10	Male	Female	MC	52	7	3	4	0	0	0
	White	Hispanic	MC	52	4	1	3	0	0	0
	White	Native American	MC	52	6	0	6	0	0	0
	No Disability	Disability	MC	52	4	1	3	1	0	1
	Not Low Income	Low Income	MC	52	1	0	1	0	0	0
	Male	Female	OR	2	2	2	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	2	0	2	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0

Table F-3. 2009–10 MontCAS: Number of Items Classified as "Low" or "High" DIF, Overall and by Group Favored – Science

Grade	Reference group	Focal group	Item type	Number of items	Number "low"			Number "high"		
					Total	Favoring reference	Favoring focal	Total	Favoring reference	Favoring focal
4	Male	Female	MC	53	5	0	5	0	0	0
	White	Hispanic	MC	53	4	0	4	0	0	0
	White	Native American	MC	53	3	0	3	0	0	0
	No Disability	Disability	MC	53	9	3	6	0	0	0
	Not Low Income	Low Income	MC	53	0	0	0	0	0	0
	Male	Female	OR	2	1	1	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	1	0	1	0	0	0
	No Disability	Disability	OR	2	0	0	0	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
8	Male	Female	MC	53	12	4	8	2	0	2
	White	Hispanic	MC	53	5	3	2	1	0	1
	White	Native American	MC	53	4	0	4	0	0	0
	No Disability	Disability	MC	53	8	2	6	1	0	1
	Not Low Income	Low Income	MC	53	0	0	0	0	0	0
	Male	Female	OR	2	1	1	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	1	0	1	0	0	0
	No Disability	Disability	OR	2	1	0	1	0	0	0
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0
10	Male	Female	MC	53	12	4	8	1	0	1
	White	Hispanic	MC	53	7	3	4	0	0	0
	White	Native American	MC	53	5	0	5	0	0	0
	No Disability	Disability	MC	53	14	3	11	2	2	0
	Not Low Income	Low Income	MC	53	0	0	0	0	0	0
	Male	Female	OR	2	1	1	0	0	0	0
	White	Hispanic	OR	2	0	0	0	0	0	0
	White	Native American	OR	2	0	0	0	0	0	0
	No Disability	Disability	OR	2	0	0	0	1	0	1
	Not Low Income	Low Income	OR	2	0	0	0	0	0	0

Appendix G—Item Response Theory Calibration Results

Table G-1. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 3

IREF	*Parameters*		
	a	*b*	*c*
60919	1.00	-1.45	0.00
76772	1.00	-1.11	0.00
76759	1.00	-1.01	0.00
76864	1.00	-0.05	0.00
76840	1.00	0.02	0.00
60313	1.00	-1.01	0.00
76781	1.00	-0.49	0.00
76795	1.00	-1.08	0.00
76750	1.00	0.62	0.00
76895	1.00	-0.36	0.00
76911	1.00	-0.13	0.00
76988	1.00	-0.59	0.00
76751	1.00	-1.16	0.00
76884	1.00	-0.09	0.00
76904	1.00	-0.94	0.00
76782	1.00	-0.17	0.00
76979	1.00	-0.26	0.00
76855	1.00	-0.53	0.00
76756	1.00	-1.92	0.00
60918	1.00	-1.29	0.00
76917	1.00	-1.40	0.00
60974	1.00	-0.79	0.00
76879	1.00	0.14	0.00
76765	1.00	-1.42	0.00
76769	1.00	-0.04	0.00
76752	1.00	-1.18	0.00
76913	1.00	0.17	0.00
77003	1.00	-1.60	0.00
61065	1.00	-0.08	0.00
76859	1.00	-0.84	0.00
60952	1.00	-0.72	0.00
76774	1.00	0.40	0.00
76909	1.00	-0.75	0.00
76841	1.00	0.46	0.00
42983	1.00	-0.16	0.00
61046	1.00	-0.31	0.00
76866	1.00	-0.92	0.00
242743	1.00	-1.45	0.00
76842	1.00	-1.22	0.00
60278	1.00	-0.91	0.00
76906	1.00	-0.45	0.00
76784	1.00	0.19	0.00
76915	1.00	-0.38	0.00
77019	1.00	-0.81	0.00
76971	1.00	-1.12	0.00
76843	1.00	0.05	0.00
43090	1.00	-1.01	0.00
77006	1.00	0.04	0.00
76777	1.00	-0.22	0.00
43136	1.00	-0.70	0.00
43105	1.00	-0.40	0.00
77011	1.00	-0.31	0.00
77027	1.00	0.49	0.00

continued

IREF	Parameters		
	a	b	c
76836	1.00	-0.58	0.00
42962	1.00	-0.67	0.00
42994	1.00	-0.42	0.00
59293	1.00	-0.80	0.00
59294	1.00	-0.83	0.00

Table G-2. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 3

IREF	Parameters					
	a	b	D1	D2	D3	D4
43261	1.00	-0.42	-1.38	0.12	0.71	0.55
76899	1.00	0.58	-1.46	-0.07	0.55	0.99

Table G-3. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 4

IREF	Parameters		
	a	b	c
76972	1.00	-1.57	0.00
242873	1.00	-0.14	0.00
62320	1.00	-0.87	0.00
76823	1.00	-0.02	0.00
76812	1.00	0.04	0.00
76961	1.00	-0.26	0.00
76892	1.00	0.18	0.00
76939	1.00	0.28	0.00
43173	1.00	-0.42	0.00
76788	1.00	0.53	0.00
76941	1.00	-0.30	0.00
244388	1.00	-0.27	0.00
43312	1.00	-0.78	0.00
43296	1.00	0.92	0.00
62363	1.00	-0.49	0.00
242880	1.00	0.07	0.00
248102	1.00	0.33	0.00
76969	1.00	-0.60	0.00
244321	1.00	-1.11	0.00
76959	1.00	-0.60	0.00
76888	1.00	-0.11	0.00
43332	1.00	-0.03	0.00
76834	1.00	0.73	0.00
43314	1.00	-0.52	0.00
62214	1.00	-0.86	0.00
76844	1.00	0.39	0.00
77022	1.00	-0.20	0.00
76794	1.00	-0.06	0.00
76935	1.00	-0.54	0.00
242889	1.00	-0.04	0.00
43330	1.00	-1.21	0.00
62401	1.00	-0.47	0.00
76924	1.00	0.24	0.00
242978	1.00	0.24	0.00
76830	1.00	-0.59	0.00
61804	1.00	0.56	0.00
61829	1.00	-0.14	0.00
76948	1.00	-0.63	0.00

continued

IREF	Parameters		
	a	*b*	*c*
76819	1.00	-0.31	0.00
77050	1.00	-0.08	0.00
76827	1.00	0.07	0.00
43298	1.00	0.15	0.00
76995	1.00	-0.51	0.00
76856	1.00	-0.27	0.00
43369	1.00	-0.07	0.00
62389	1.00	-0.29	0.00
76965	1.00	0.12	0.00
76933	1.00	0.15	0.00
243037	1.00	-0.41	0.00
61811	1.00	0.99	0.00
43167	1.00	-0.02	0.00
43276	1.00	-0.57	0.00
62381	1.00	-0.71	0.00
76963	1.00	0.47	0.00
43320	1.00	-0.73	0.00
243174	1.00	-0.40	0.00
61780	1.00	0.18	0.00
76763	1.00	0.31	0.00

Table G-4. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 4

IREF	Parameters					
	a	*b*	*D1*	*D2*	*D3*	*D4*
77063	1.00	0.34	-0.92	0.41	-0.42	0.93
62486	1.00	0.31	0.19	0.10	-0.07	-0.22

Table G-5. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 5

IREF	Parameters		
	a	*b*	*c*
60551	1.00	-1.13	0.00
77210	1.00	-0.33	0.00
77270	1.00	-0.42	0.00
77314	1.00	0.05	0.00
60417	1.00	-0.43	0.00
77325	1.00	-0.92	0.00
77247	1.00	-0.42	0.00
77259	1.00	-0.99	0.00
77179	1.00	-0.34	0.00
77274	1.00	-0.83	0.00
77193	1.00	-0.86	0.00
43514	1.00	0.29	0.00
77228	1.00	0.07	0.00
43560	1.00	-0.10	0.00
77245	1.00	-0.15	0.00
77205	1.00	-0.77	0.00
77310	1.00	-0.06	0.00
77220	1.00	-0.29	0.00
77200	1.00	-0.80	0.00
77163	1.00	-0.41	0.00
43477	1.00	-0.36	0.00
77318	1.00	0.31	0.00
60370	1.00	-0.31	0.00
77257	1.00	-0.40	0.00

continued

IREF	Parameters		
	a	*b*	*c*
77204	1.00	-0.25	0.00
43471	1.00	-0.49	0.00
77208	1.00	0.34	0.00
77181	1.00	-0.33	0.00
77191	1.00	-0.58	0.00
59800	1.00	-0.32	0.00
77243	1.00	-0.06	0.00
43409	1.00	0.51	0.00
77279	1.00	-0.63	0.00
77207	1.00	-0.10	0.00
59986	1.00	-0.56	0.00
77330	1.00	0.53	0.00
43413	1.00	-0.07	0.00
43516	1.00	-0.93	0.00
77219	1.00	0.49	0.00
77321	1.00	0.58	0.00
59856	1.00	-0.33	0.00
77186	1.00	-0.38	0.00
60504	1.00	-0.69	0.00
77265	1.00	0.42	0.00
60371	1.00	-0.98	0.00
60971	1.00	-0.16	0.00
77217	1.00	-0.60	0.00
77198	1.00	0.36	0.00
77255	1.00	-0.19	0.00
242955	1.00	0.21	0.00
77177	1.00	-0.87	0.00
77282	1.00	-0.65	0.00
43585	1.00	-0.32	0.00
60911	1.00	-0.35	0.00
43421	1.00	-0.72	0.00
77298	1.00	0.58	0.00
62025	1.00	-0.38	0.00
77295	1.00	-1.04	0.00

Table G-6. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 5

IREF	Parameters					
	a	*b*	*D1*	*D2*	*D3*	*D4*
61052	1.00	0.41	-0.86	0.42	-0.09	0.52
242957	1.00	-0.33	-0.01	-0.61	0.46	0.16

Table G-7. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 6

IREF	Parameters		
	a	*b*	*c*
77377	1.00	-1.93	0.00
77320	1.00	-0.91	0.00
77376	1.00	-0.23	0.00
77340	1.00	-0.12	0.00
77323	1.00	-0.56	0.00
60885	1.00	-0.30	0.00
60901	1.00	-0.04	0.00
77378	1.00	-0.10	0.00
77317	1.00	0.12	0.00
43912	1.00	0.05	0.00

continued

IREF	Parameters		
	a	*b*	*c*
77373	1.00	-0.05	0.00
43863	1.00	0.24	0.00
43921	1.00	0.05	0.00
77313	1.00	0.36	0.00
44004	1.00	-0.70	0.00
62994	1.00	0.25	0.00
77522	1.00	-0.27	0.00
243125	1.00	-0.52	0.00
77630	1.00	-0.09	0.00
43992	1.00	-0.05	0.00
77614	1.00	-0.71	0.00
77339	1.00	0.08	0.00
77476	1.00	-0.67	0.00
77553	1.00	-0.06	0.00
77451	1.00	0.15	0.00
77398	1.00	0.58	0.00
77608	1.00	0.39	0.00
61156	1.00	-0.03	0.00
77515	1.00	0.16	0.00
77509	1.00	-0.26	0.00
77538	1.00	-0.60	0.00
62050	1.00	0.27	0.00
77414	1.00	-0.63	0.00
77542	1.00	0.52	0.00
43995	1.00	-0.84	0.00
77445	1.00	-1.84	0.00
77621	1.00	-0.57	0.00
43956	1.00	-0.22	0.00
77502	1.00	-0.23	0.00
77582	1.00	-0.35	0.00
62060	1.00	-0.94	0.00
77380	1.00	-0.08	0.00
43944	1.00	-0.20	0.00
242549	1.00	0.46	0.00
77449	1.00	0.14	0.00
43949	1.00	0.18	0.00
62029	1.00	-0.24	0.00
77625	1.00	-1.22	0.00
77633	1.00	-0.23	0.00
43991	1.00	0.13	0.00
43963	1.00	-0.53	0.00
243201	1.00	0.20	0.00
77455	1.00	-0.41	0.00
77497	1.00	0.38	0.00
77517	1.00	-0.20	0.00
63005	1.00	0.18	0.00
44088	1.00	-1.04	0.00
77642	1.00	0.11	0.00

Table G-8. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 6

IREF	Parameters					
	a	*b*	*D1*	*D2*	*D3*	*D4*
77649	1.00	0.59	0.03	-0.51	0.85	-0.37
44048	1.00	0.38	-0.87	0.02	0.00	0.84

Table G-9. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 7

IREF	*Parameters*		
	a	*b*	*c*
61206	1.00	-0.67	0.00
86297	1.00	-0.40	0.00
61204	1.00	0.16	0.00
86366	1.00	-0.04	0.00
86305	1.00	0.38	0.00
61205	1.00	0.50	0.00
43836	1.00	-0.27	0.00
62948	1.00	0.54	0.00
86280	1.00	-0.76	0.00
86296	1.00	-0.08	0.00
86300	1.00	0.58	0.00
86302	1.00	0.01	0.00
61228	1.00	0.27	0.00
86295	1.00	-0.46	0.00
61742	1.00	0.02	0.00
86431	1.00	0.33	0.00
43787	1.00	-0.30	0.00
86438	1.00	-1.74	0.00
86382	1.00	-0.72	0.00
86615	1.00	0.40	0.00
86555	1.00	-0.58	0.00
43721	1.00	-0.01	0.00
86549	1.00	0.04	0.00
43782	1.00	0.14	0.00
86455	1.00	0.03	0.00
86635	1.00	-0.97	0.00
86395	1.00	0.27	0.00
61745	1.00	-0.29	0.00
86374	1.00	-0.24	0.00
61279	1.00	-1.20	0.00
86631	1.00	0.25	0.00
61195	1.00	0.47	0.00
43705	1.00	2.10	0.00
86381	1.00	0.43	0.00
86379	1.00	-0.31	0.00
86535	1.00	-0.84	0.00
86369	1.00	-0.31	0.00
86448	1.00	0.08	0.00
86473	1.00	-0.27	0.00
86692	1.00	-0.19	0.00
61766	1.00	-1.14	0.00
86458	1.00	0.14	0.00
86650	1.00	-0.52	0.00
61250	1.00	-0.52	0.00
86683	1.00	-0.66	0.00
86622	1.00	-0.09	0.00
86311	1.00	0.08	0.00
86313	1.00	-0.16	0.00
86453	1.00	0.49	0.00
86591	1.00	-1.10	0.00
86482	1.00	-0.70	0.00
86568	1.00	0.10	0.00
86689	1.00	-0.30	0.00
86675	1.00	0.28	0.00
43700	1.00	-0.04	0.00
86348	1.00	-0.23	0.00

continued

IREF	Parameters		
	a	b	c
43799	1.00	-0.40	0.00
86349	1.00	0.11	0.00

Table G-10. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 7

IREF	Parameters					
	a	b	D1	D2	D3	D4
158633	1.00	0.53	-0.97	0.13	0.39	0.44
43922	1.00	0.17	0.29	-0.91	0.60	0.02

Table G-11. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 8

IREF	Parameters		
	a	b	c
87598	1.00	-1.67	0.00
63025	1.00	0.27	0.00
87593	1.00	-0.64	0.00
87658	1.00	0.08	0.00
87527	1.00	-0.34	0.00
87583	1.00	-0.37	0.00
87808	1.00	0.51	0.00
87802	1.00	-0.26	0.00
87661	1.00	0.77	0.00
87580	1.00	-0.41	0.00
44214	1.00	0.15	0.00
87606	1.00	-0.01	0.00
44176	1.00	0.58	0.00
44626	1.00	-1.11	0.00
88363	1.00	0.39	0.00
88177	1.00	-0.57	0.00
44160	1.00	0.13	0.00
88019	1.00	-0.17	0.00
88864	1.00	-0.27	0.00
87623	1.00	0.51	0.00
44137	1.00	0.04	0.00
44141	1.00	-0.81	0.00
88838	1.00	-0.33	0.00
63148	1.00	0.66	0.00
63219	1.00	0.02	0.00
88174	1.00	-0.17	0.00
88189	1.00	0.74	0.00
63144	1.00	0.33	0.00
62943	1.00	0.04	0.00
44642	1.00	1.39	0.00
88338	1.00	0.00	0.00
44662	1.00	0.19	0.00
44220	1.00	-0.11	0.00
88216	1.00	-0.42	0.00
63038	1.00	-0.55	0.00
87799	1.00	0.03	0.00
243793	1.00	0.34	0.00
244538	1.00	-1.30	0.00
44205	1.00	-0.73	0.00
63252	1.00	0.04	0.00
63203	1.00	0.32	0.00
44123	1.00	0.19	0.00

continued

IREF	Parameters		
	a	*b*	*c*
63115	1.00	0.30	0.00
44210	1.00	-0.26	0.00
44648	1.00	-0.01	0.00
88086	1.00	0.06	0.00
88848	1.00	-0.13	0.00
63287	1.00	-0.38	0.00
88263	1.00	-0.06	0.00
63106	1.00	-0.87	0.00
63256	1.00	-0.43	0.00
86642	1.00	0.71	0.00
88381	1.00	-0.17	0.00
88352	1.00	-0.37	0.00
44199	1.00	0.04	0.00
87834	1.00	0.12	0.00
87841	1.00	0.16	0.00

Table G-12. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 8

IREF	Parameters					
	a	*b*	*D1*	*D2*	*D3*	*D4*
44267	1.00	0.33	-0.10	-0.47	0.16	0.41

Table G-13. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Mathematics Grade 10

IREF	Parameters		
	a	*b*	*c*
59397	1.00	-0.53	0.00
77480	1.00	0.07	0.00
77571	1.00	0.42	0.00
61319	1.00	0.32	0.00
77612	1.00	-0.15	0.00
61298	1.00	0.54	0.00
77623	1.00	-0.06	0.00
77618	1.00	-0.71	0.00
77570	1.00	0.11	0.00
59377	1.00	-0.07	0.00
77596	1.00	-0.09	0.00
43613	1.00	0.81	0.00
59365	1.00	0.54	0.00
243091	1.00	0.05	0.00
77368	1.00	-0.15	0.00
43969	1.00	-0.98	0.00
62368	1.00	-0.14	0.00
77507	1.00	0.24	0.00
77352	1.00	0.97	0.00
77371	1.00	-0.75	0.00
77485	1.00	-0.13	0.00
243153	1.00	0.30	0.00
77392	1.00	-0.02	0.00
43743	1.00	0.83	0.00
243118	1.00	0.34	0.00
77382	1.00	0.41	0.00
43822	1.00	-0.29	0.00
43917	1.00	0.06	0.00
242987	1.00	-0.68	0.00
62292	1.00	0.34	0.00

continued

IREF	Parameters		
	a	b	c
77384	1.00	0.56	0.00
77432	1.00	-0.29	0.00
77503	1.00	0.60	0.00
62202	1.00	-0.10	0.00
243121	1.00	-0.19	0.00
62177	1.00	-1.09	0.00
77484	1.00	-0.49	0.00
77370	1.00	-0.06	0.00
77354	1.00	-0.22	0.00
77415	1.00	-0.01	0.00
77562	1.00	0.47	0.00
61265	1.00	0.37	0.00
77428	1.00	-0.27	0.00
77561	1.00	-0.07	0.00
62333	1.00	-0.23	0.00
77394	1.00	0.35	0.00
61324	1.00	-1.31	0.00
248846	1.00	0.82	0.00
77551	1.00	0.03	0.00
77619	1.00	0.09	0.00
61281	1.00	-0.56	0.00
62286	1.00	0.47	0.00
243134	1.00	0.24	0.00
240995	1.00	0.60	0.00
61312	1.00	-0.21	0.00
43899	1.00	0.41	0.00
243139	1.00	0.52	0.00
77635	1.00	-0.01	0.00

Table G-14. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Mathematics Grade 10

IREF	Parameters					
	a	b	D1	D2	D3	D4
160523	1.00	0.52	-0.34	-0.39	0.88	-0.14
159645	1.00	0.93	-0.84	-0.64	0.65	0.83

Table G-15. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 3

IREF	Parameters		
	a	b	c
68808	1.00	-0.86	0.00
68809	1.00	-0.41	0.00
68811	1.00	-1.45	0.00
68810	1.00	0.11	0.00
68812	1.00	-0.02	0.00
68814	1.00	-0.02	0.00
68818	1.00	-0.26	0.00
92739	1.00	-0.81	0.00
92742	1.00	-0.40	0.00
92743	1.00	-0.33	0.00
92745	1.00	0.04	0.00
92746	1.00	-1.14	0.00
92748	1.00	-1.04	0.00
92749	1.00	-0.79	0.00
92750	1.00	-0.65	0.00
92751	1.00	-0.59	0.00

continued

IREF	Parameters		
	a	*b*	*c*
92758	1.00	-0.79	0.00
92752	1.00	-0.23	0.00
92755	1.00	-0.84	0.00
92786	1.00	-0.09	0.00
92789	1.00	-1.32	0.00
92791	1.00	-1.20	0.00
92792	1.00	-0.69	0.00
92794	1.00	-0.55	0.00
92797	1.00	-0.60	0.00
92795	1.00	-0.16	0.00
92670	1.00	-1.37	0.00
92673	1.00	-0.16	0.00
92674	1.00	0.07	0.00
92675	1.00	-1.01	0.00
92677	1.00	-0.20	0.00
92695	1.00	-0.81	0.00
92696	1.00	-0.18	0.00
92798	1.00	-0.95	0.00
92800	1.00	-0.65	0.00
157468	1.00	-0.20	0.00
92802	1.00	0.02	0.00
92803	1.00	-0.60	0.00
92808	1.00	-0.41	0.00
157469	1.00	0.10	0.00
92763	1.00	-0.50	0.00
92765	1.00	-0.06	0.00
92767	1.00	-0.84	0.00
92771	1.00	-0.09	0.00
92768	1.00	-0.82	0.00
92766	1.00	-0.37	0.00
92773	1.00	-0.51	0.00
92775	1.00	-1.10	0.00
92777	1.00	-0.09	0.00
92781	1.00	-0.12	0.00
92779	1.00	-0.06	0.00
92778	1.00	-0.26	0.00

Table G-16. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 3

IREF	Parameters					
	a	*b*	*D1*	*D2*	*D3*	*D4*
92761	1.00	0.66	-1.46	-0.79	0.64	1.60
92783	1.00	0.70	-0.80	-0.64	0.34	1.10

Table G-17. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 4

IREF	Parameters		
	a	*b*	*c*
93939	1.00	0.06	0.00
157543	1.00	-0.73	0.00
93944	1.00	-0.31	0.00
93946	1.00	-0.59	0.00
93948	1.00	-0.20	0.00
93951	1.00	-0.43	0.00
157544	1.00	0.09	0.00
67330	1.00	-0.28	0.00

continued

IREF	Parameters		
	a	b	c
67333	1.00	-0.01	0.00
67334	1.00	0.18	0.00
67365	1.00	-0.14	0.00
67350	1.00	-0.40	0.00
67354	1.00	0.18	0.00
67346	1.00	-0.35	0.00
67359	1.00	0.12	0.00
67367	1.00	0.13	0.00
67371	1.00	-0.46	0.00
67374	1.00	-0.23	0.00
67368	1.00	-0.08	0.00
93837	1.00	-0.48	0.00
93838	1.00	-0.06	0.00
93846	1.00	0.02	0.00
93876	1.00	-0.56	0.00
93881	1.00	-0.21	0.00
93895	1.00	-0.63	0.00
93888	1.00	0.34	0.00
93842	1.00	0.23	0.00
93855	1.00	-1.07	0.00
93857	1.00	-0.20	0.00
93862	1.00	-0.38	0.00
93863	1.00	-0.44	0.00
93849	1.00	-0.18	0.00
93874	1.00	-0.17	0.00
94002	1.00	-0.37	0.00
94004	1.00	-0.26	0.00
157546	1.00	-0.17	0.00
94016	1.00	-0.02	0.00
94015	1.00	-0.57	0.00
94027	1.00	-0.04	0.00
157548	1.00	0.28	0.00
94048	1.00	-0.28	0.00
94050	1.00	-0.25	0.00
94072	1.00	-0.11	0.00
94083	1.00	0.35	0.00
94079	1.00	-0.84	0.00
94108	1.00	-0.41	0.00
94092	1.00	-0.70	0.00
94095	1.00	-0.29	0.00
94077	1.00	-0.50	0.00
94111	1.00	-0.41	0.00
94120	1.00	-0.29	0.00
94116	1.00	0.55	0.00

Table G-18. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 4

IREF	Parameters					
	a	b	D1	D2	D3	D4
67382	1.00	0.84	-1.08	-0.51	0.27	1.32
94139	1.00	0.71	-1.42	-1.09	0.82	1.69

Table G-19. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 5

IREF	*Parameters*		
	a	*b*	*c*
65312	1.00	-0.66	0.00
65317	1.00	-0.43	0.00
65383	1.00	-0.64	0.00
65368	1.00	-0.29	0.00
69235	1.00	-0.73	0.00
65379	1.00	-0.89	0.00
65387	1.00	0.21	0.00
93601	1.00	-0.64	0.00
93611	1.00	-0.23	0.00
93612	1.00	-0.12	0.00
93615	1.00	-0.84	0.00
93616	1.00	-0.29	0.00
93626	1.00	-0.14	0.00
93631	1.00	-0.82	0.00
93623	1.00	-0.48	0.00
93628	1.00	-0.37	0.00
93635	1.00	-0.35	0.00
93639	1.00	-1.20	0.00
93638	1.00	-0.22	0.00
93353	1.00	-0.03	0.00
93366	1.00	-0.43	0.00
93375	1.00	-1.13	0.00
93378	1.00	-0.39	0.00
93381	1.00	-0.86	0.00
93389	1.00	-0.17	0.00
93385	1.00	0.08	0.00
93529	1.00	-0.65	0.00
93510	1.00	-0.25	0.00
93533	1.00	-0.14	0.00
93524	1.00	-0.44	0.00
93520	1.00	-0.27	0.00
93536	1.00	-0.50	0.00
93537	1.00	-0.05	0.00
93700	1.00	-1.05	0.00
93695	1.00	-0.18	0.00
93698	1.00	-0.71	0.00
93705	1.00	0.04	0.00
93704	1.00	-0.59	0.00
93709	1.00	0.04	0.00
93711	1.00	-0.80	0.00
93448	1.00	-0.58	0.00
93414	1.00	-0.75	0.00
93415	1.00	-0.60	0.00
93416	1.00	-0.64	0.00
93419	1.00	-1.19	0.00
93420	1.00	-0.65	0.00
93431	1.00	-0.23	0.00
93421	1.00	-0.75	0.00
93428	1.00	-0.86	0.00
93446	1.00	0.14	0.00
93444	1.00	-0.45	0.00
93451	1.00	0.09	0.00

Table G-20. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 5

IREF	*Parameters*					
	a	*b*	*D1*	*D2*	*D3*	*D4*
93668	1.00	0.37	-0.74	-0.70	0.04	1.40
93462	1.00	0.17	-1.60	-0.96	0.69	1.87

Table G-21. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 6

IREF	*Parameters*		
	a	*b*	*c*
95410	1.00	-0.33	0.00
95421	1.00	-0.95	0.00
95409	1.00	-0.67	0.00
95445	1.00	0.01	0.00
95431	1.00	-0.20	0.00
95435	1.00	-0.15	0.00
95450	1.00	-0.62	0.00
95305	1.00	-0.29	0.00
95335	1.00	-0.33	0.00
95330	1.00	-0.38	0.00
95351	1.00	-0.45	0.00
95358	1.00	-0.78	0.00
95363	1.00	-0.92	0.00
95369	1.00	-0.62	0.00
95371	1.00	-0.90	0.00
95375	1.00	-0.40	0.00
95386	1.00	0.09	0.00
95381	1.00	-0.07	0.00
95393	1.00	-0.12	0.00
95077	1.00	0.26	0.00
95088	1.00	-0.94	0.00
95092	1.00	-0.85	0.00
95101	1.00	-0.71	0.00
95132	1.00	-0.21	0.00
95105	1.00	0.21	0.00
95115	1.00	-0.34	0.00
95348	1.00	-0.39	0.00
95342	1.00	-0.73	0.00
95353	1.00	-0.33	0.00
95368	1.00	-1.04	0.00
97773	1.00	-0.55	0.00
95387	1.00	-0.34	0.00
95398	1.00	-0.56	0.00
95202	1.00	-0.21	0.00
95183	1.00	0.35	0.00
95218	1.00	-0.36	0.00
95228	1.00	-0.64	0.00
95231	1.00	-0.51	0.00
95289	1.00	0.03	0.00
95299	1.00	0.14	0.00
95033	1.00	-0.83	0.00
95036	1.00	-0.77	0.00
95041	1.00	0.17	0.00
95045	1.00	-0.80	0.00
95085	1.00	-1.21	0.00
95089	1.00	-0.26	0.00
95093	1.00	-0.73	0.00
95121	1.00	0.17	0.00
95110	1.00	-0.58	0.00

continued

IREF	Parameters		
	a	b	c
95114	1.00	-0.63	0.00
95145	1.00	-0.26	0.00
95157	1.00	-0.72	0.00

Table G-22. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 6

IREF	Parameters					
	a	b	D1	D2	D3	D4
95397	1.00	0.17	-0.97	-0.61	0.49	1.08
95171	1.00	0.22	-1.16	-0.92	0.55	1.53

Table G-23. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 7

IREF	Parameters		
	a	b	c
41859	1.00	-0.09	0.00
157578	1.00	0.04	0.00
41860	1.00	-0.41	0.00
41867	1.00	-0.81	0.00
41864	1.00	-0.85	0.00
41866	1.00	-0.41	0.00
41868	1.00	-0.84	0.00
92567	1.00	-0.65	0.00
157579	1.00	-0.71	0.00
92583	1.00	-0.18	0.00
92588	1.00	-0.33	0.00
92589	1.00	-0.35	0.00
92591	1.00	-1.00	0.00
92593	1.00	-0.56	0.00
92596	1.00	-0.13	0.00
92600	1.00	-0.85	0.00
92605	1.00	-0.49	0.00
92606	1.00	0.00	0.00
92608	1.00	-0.70	0.00
68610	1.00	-0.99	0.00
68611	1.00	-0.12	0.00
68612	1.00	-0.49	0.00
68616	1.00	-0.59	0.00
68613	1.00	-0.36	0.00
68614	1.00	0.12	0.00
68620	1.00	-0.95	0.00
92341	1.00	-0.20	0.00
92342	1.00	-0.33	0.00
92343	1.00	-0.53	0.00
92345	1.00	-0.48	0.00
92348	1.00	-0.21	0.00
92350	1.00	-0.51	0.00
92347	1.00	-0.30	0.00
68493	1.00	-1.05	0.00
68495	1.00	-0.10	0.00
68497	1.00	-0.59	0.00
68510	1.00	-0.83	0.00
68507	1.00	-0.98	0.00
68498	1.00	-0.84	0.00
68514	1.00	-0.48	0.00
92531	1.00	-0.10	0.00

continued

IREF	Parameters		
	a	b	c
92535	1.00	-0.49	0.00
92536	1.00	-0.24	0.00
92540	1.00	-0.02	0.00
92541	1.00	-0.17	0.00
92543	1.00	-0.52	0.00
92545	1.00	-0.06	0.00
92549	1.00	-0.94	0.00
92554	1.00	-0.87	0.00
92555	1.00	-0.36	0.00
92558	1.00	0.01	0.00
92559	1.00	0.07	0.00

Table G-24. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 7

IREF	Parameters					
	a	b	D1	D2	D3	D4
92611	1.00	0.06	-1.36	-0.64	0.42	1.58
92562	1.00	0.11	-1.36	-0.59	0.27	1.68

Table G-25. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 8

IREF	Parameters		
	a	b	c
67937	1.00	-0.11	0.00
67938	1.00	-0.44	0.00
67944	1.00	-0.45	0.00
67948	1.00	-0.05	0.00
67953	1.00	0.28	0.00
67952	1.00	-1.02	0.00
67966	1.00	-0.65	0.00
95838	1.00	-0.17	0.00
95843	1.00	-0.28	0.00
95844	1.00	0.31	0.00
95845	1.00	-0.29	0.00
95847	1.00	-0.19	0.00
95851	1.00	-0.67	0.00
95853	1.00	-0.45	0.00
95855	1.00	-0.08	0.00
95856	1.00	-0.56	0.00
95863	1.00	-0.57	0.00
95867	1.00	-0.58	0.00
95866	1.00	-0.07	0.00
68698	1.00	-0.10	0.00
68699	1.00	-0.26	0.00
68702	1.00	-0.04	0.00
68714	1.00	-0.36	0.00
68725	1.00	-0.42	0.00
68724	1.00	-0.57	0.00
68726	1.00	-0.71	0.00
95688	1.00	-0.20	0.00
95691	1.00	-0.37	0.00
95700	1.00	-0.02	0.00
95703	1.00	-0.19	0.00
157551	1.00	0.28	0.00
95706	1.00	-0.76	0.00
95708	1.00	-0.14	0.00

continued

IREF	Parameters		
	a	b	c
95604	1.00	-0.48	0.00
95627	1.00	-0.47	0.00
95637	1.00	-0.18	0.00
95644	1.00	-0.07	0.00
95647	1.00	-0.16	0.00
95649	1.00	0.29	0.00
95656	1.00	-0.35	0.00
68470	1.00	-0.06	0.00
68473	1.00	-0.34	0.00
68476	1.00	-0.45	0.00
68478	1.00	-0.46	0.00
68475	1.00	-0.37	0.00
68480	1.00	-0.05	0.00
68487	1.00	-0.05	0.00
68500	1.00	-0.50	0.00
68499	1.00	-0.42	0.00
68504	1.00	-0.23	0.00
68508	1.00	0.07	0.00
68501	1.00	-0.22	0.00

Table G-26. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 8

IREF	Parameters					
	a	b	D1	D2	D3	D4
95869	1.00	0.37	-1.32	-0.69	0.48	1.52
68511	1.00	0.06	-1.40	-0.66	0.18	1.87

Table G-27. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Reading Grade 10

IREF	Parameters		
	a	b	c
67599	1.00	-0.89	0.00
67687	1.00	-0.49	0.00
67720	1.00	0.22	0.00
67741	1.00	-0.06	0.00
67750	1.00	-0.48	0.00
67753	1.00	-0.61	0.00
67757	1.00	-0.19	0.00
94842	1.00	-0.02	0.00
94843	1.00	-0.65	0.00
94848	1.00	-0.31	0.00
94852	1.00	-0.88	0.00
94854	1.00	-0.83	0.00
94857	1.00	-0.95	0.00
94861	1.00	-0.29	0.00
94874	1.00	-0.19	0.00
94877	1.00	-0.40	0.00
94879	1.00	0.04	0.00
94863	1.00	-0.16	0.00
94882	1.00	0.15	0.00
95338	1.00	-0.49	0.00
95340	1.00	-0.39	0.00
95361	1.00	-0.33	0.00
95367	1.00	-0.24	0.00
95374	1.00	0.03	0.00
95377	1.00	-0.94	0.00

continued

IREF	Parameters		
	a	b	c
95391	1.00	-0.02	0.00
95026	1.00	-0.45	0.00
95030	1.00	-0.64	0.00
95138	1.00	-0.10	0.00
95164	1.00	-0.39	0.00
95187	1.00	-0.06	0.00
95154	1.00	-0.56	0.00
95207	1.00	-0.97	0.00
95216	1.00	-0.47	0.00
95273	1.00	-0.10	0.00
95234	1.00	-0.31	0.00
95279	1.00	0.02	0.00
95285	1.00	-0.03	0.00
95290	1.00	0.18	0.00
95293	1.00	-0.56	0.00
94889	1.00	-0.65	0.00
94890	1.00	0.13	0.00
94892	1.00	-0.13	0.00
94894	1.00	0.00	0.00
94898	1.00	0.03	0.00
94903	1.00	-0.33	0.00
94912	1.00	-0.19	0.00
94924	1.00	0.10	0.00
94929	1.00	-0.41	0.00
94931	1.00	-0.63	0.00
94941	1.00	-0.62	0.00
94943	1.00	-0.57	0.00

Table G-28. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Reading Grade 10

IREF	Parameters					
	a	b	D1	D2	D3	D4
94887	1.00	0.26	-1.30	-0.65	0.57	1.38
94955	1.00	0.50	-0.94	-0.44	0.24	1.13

Table G-29. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Science Grade 4

IREF	Parameters		
	a	b	c
39067	1.00	-0.07	0.00
75824	1.00	-0.74	0.00
75690	1.00	0.04	0.00
39242	1.00	-0.48	0.00
39314	1.00	-0.04	0.00
57874	1.00	-1.08	0.00
53659	1.00	-1.23	0.00
75923	1.00	0.01	0.00
75835	1.00	-0.30	0.00
39229	1.00	-0.21	0.00
39086	1.00	-0.20	0.00
76403	1.00	-0.97	0.00
39336	1.00	0.37	0.00
42786	1.00	0.07	0.00
76285	1.00	-0.85	0.00
39184	1.00	-0.43	0.00
75889	1.00	0.07	0.00

continued

IREF	Parameters		
	a	*b*	*c*
75902	1.00	-0.63	0.00
39119	1.00	-0.37	0.00
76394	1.00	-0.59	0.00
75784	1.00	-0.07	0.00
75910	1.00	-0.97	0.00
39127	1.00	-0.01	0.00
39257	1.00	-0.38	0.00
39238	1.00	-0.31	0.00
76296	1.00	-0.67	0.00
75717	1.00	-0.46	0.00
75418	1.00	-0.51	0.00
75737	1.00	-0.84	0.00
75788	1.00	-1.09	0.00
75833	1.00	-1.09	0.00
75741	1.00	-0.38	0.00
75912	1.00	-0.56	0.00
75908	1.00	-0.60	0.00
75702	1.00	-1.18	0.00
55629	1.00	0.29	0.00
75421	1.00	-0.11	0.00
75887	1.00	-0.71	0.00
75694	1.00	-0.46	0.00
53393	1.00	-0.37	0.00
75828	1.00	-0.53	0.00
75801	1.00	-0.10	0.00
75782	1.00	0.19	0.00
75901	1.00	-0.38	0.00
75895	1.00	-0.20	0.00
75899	1.00	-0.53	0.00
75408	1.00	0.02	0.00
55464	1.00	-1.13	0.00
57860	1.00	-0.29	0.00
75517	1.00	-0.14	0.00
56340	1.00	-0.25	0.00
75423	1.00	-1.08	0.00
75752	1.00	0.12	0.00

Table G-30. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Science Grade 4

IREF	Parameters					
	a	*b*	*D1*	*D2*	*D3*	*D4*
42790	1.00	-0.33	-0.38	0.01	-0.15	0.52
75427	1.00	0.52	-0.38	-0.90	0.27	1.01

Table G-31. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Science Grade 8

IREF	Parameters		
	a	*b*	*c*
89277	1.00	-1.02	0.00
89520	1.00	-0.83	0.00
39701	1.00	-0.57	0.00
89817	1.00	-0.54	0.00
89693	1.00	-0.30	0.00
39782	1.00	0.01	0.00
39745	1.00	-0.41	0.00
89647	1.00	-0.52	0.00

continued

IREF	Parameters		
	a	*b*	*c*
89361	1.00	-0.38	0.00
89911	1.00	-0.35	0.00
89593	1.00	-0.67	0.00
56851	1.00	0.33	0.00
89582	1.00	0.55	0.00
89850	1.00	0.21	0.00
89522	1.00	-0.28	0.00
89892	1.00	-0.32	0.00
89639	1.00	-0.01	0.00
56805	1.00	-1.01	0.00
89420	1.00	-0.75	0.00
89263	1.00	-0.39	0.00
39659	1.00	0.19	0.00
56828	1.00	0.23	0.00
89498	1.00	-0.43	0.00
89778	1.00	0.13	0.00
89652	1.00	0.38	0.00
89795	1.00	-0.48	0.00
89742	1.00	0.51	0.00
89691	1.00	0.00	0.00
89457	1.00	0.19	0.00
56897	1.00	-0.42	0.00
89884	1.00	-0.23	0.00
89468	1.00	0.01	0.00
89634	1.00	0.12	0.00
89452	1.00	-0.41	0.00
89766	1.00	-0.34	0.00
38602	1.00	-0.67	0.00
89752	1.00	-0.41	0.00
54264	1.00	-0.57	0.00
89770	1.00	0.00	0.00
89444	1.00	0.37	0.00
89726	1.00	0.06	0.00
39652	1.00	0.15	0.00
56833	1.00	0.18	0.00
89849	1.00	-0.22	0.00
89508	1.00	0.16	0.00
56992	1.00	0.21	0.00
54543	1.00	-0.21	0.00
39780	1.00	-0.75	0.00
89505	1.00	0.42	0.00
54454	1.00	-0.23	0.00
89870	1.00	-0.42	0.00
89382	1.00	-0.59	0.00
89863	1.00	0.00	0.00

Table G-32. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Science Grade 8

IREF	Parameters					
	a	*b*	*D1*	*D2*	*D3*	*D4*
55106	1.00	-0.41	-0.62	-0.03	0.33	0.32
39764	1.00	0.08	0.31	-0.20	0.27	-0.38

Table G-33. 2009–10 MontCAS: IRT Parameters for Dichotomous Items – Science Grade 10

IREF	*Parameters*		
	a	*b*	*c*
40317	1.00	-0.64	0.00
55819	1.00	-0.68	0.00
75948	1.00	0.06	0.00
75876	1.00	-0.27	0.00
60856	1.00	-0.33	0.00
53812	1.00	0.04	0.00
75445	1.00	0.41	0.00
75650	1.00	-0.73	0.00
75958	1.00	0.33	0.00
75433	1.00	-0.06	0.00
53584	1.00	0.38	0.00
55620	1.00	-0.42	0.00
75440	1.00	0.12	0.00
75780	1.00	-0.35	0.00
53265	1.00	-0.55	0.00
75456	1.00	-0.12	0.00
75634	1.00	0.16	0.00
40331	1.00	-0.98	0.00
75739	1.00	-0.47	0.00
75963	1.00	-0.25	0.00
75611	1.00	-0.23	0.00
75880	1.00	-0.32	0.00
75629	1.00	-0.37	0.00
75859	1.00	-0.21	0.00
75764	1.00	-0.97	0.00
75635	1.00	-0.02	0.00
75807	1.00	-0.14	0.00
75442	1.00	-0.07	0.00
75856	1.00	-0.75	0.00
56702	1.00	-0.16	0.00
75941	1.00	-0.47	0.00
75785	1.00	-0.52	0.00
75861	1.00	0.11	0.00
75878	1.00	0.47	0.00
75706	1.00	-0.29	0.00
40285	1.00	-0.71	0.00
55696	1.00	-1.63	0.00
54221	1.00	-0.35	0.00
75873	1.00	0.03	0.00
75980	1.00	0.05	0.00
55710	1.00	-0.18	0.00
75620	1.00	0.45	0.00
75972	1.00	-0.22	0.00
75804	1.00	0.10	0.00
56698	1.00	0.13	0.00
75631	1.00	-0.10	0.00
56234	1.00	0.12	0.00
75863	1.00	-0.41	0.00
52985	1.00	0.01	0.00
75854	1.00	-0.22	0.00
75970	1.00	-0.14	0.00
75638	1.00	-0.56	0.00
75937	1.00	-0.84	0.00

Table G-34. 2009–10 MontCAS: IRT Parameters for Polytomous Items – Science Grade 10

IREF	*Parameters*					
	a	*b*	*D1*	*D2*	*D3*	*D4*
75461	1.00	0.21	-0.90	0.15	0.25	0.50
75652	1.00	0.90	-0.95	-0.33	0.40	0.88

Appendix H—TCCs AND TIFs

Figure H-1. 2009–10 MontCAS: TCC – Mathematics Grade 3

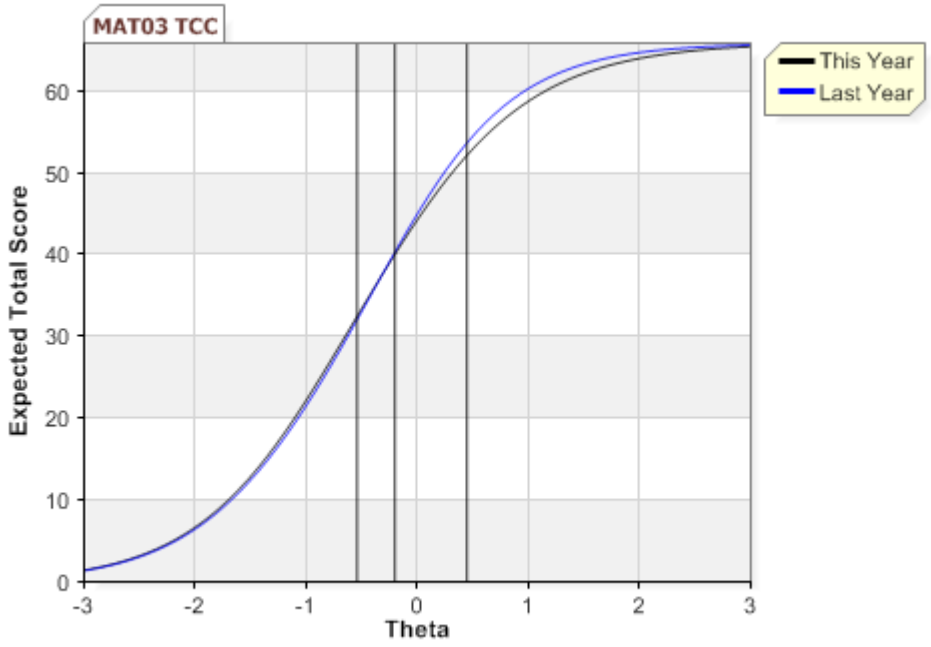


Figure H-2. 2009–10 MontCAS: TIF – Mathematics Grade 3

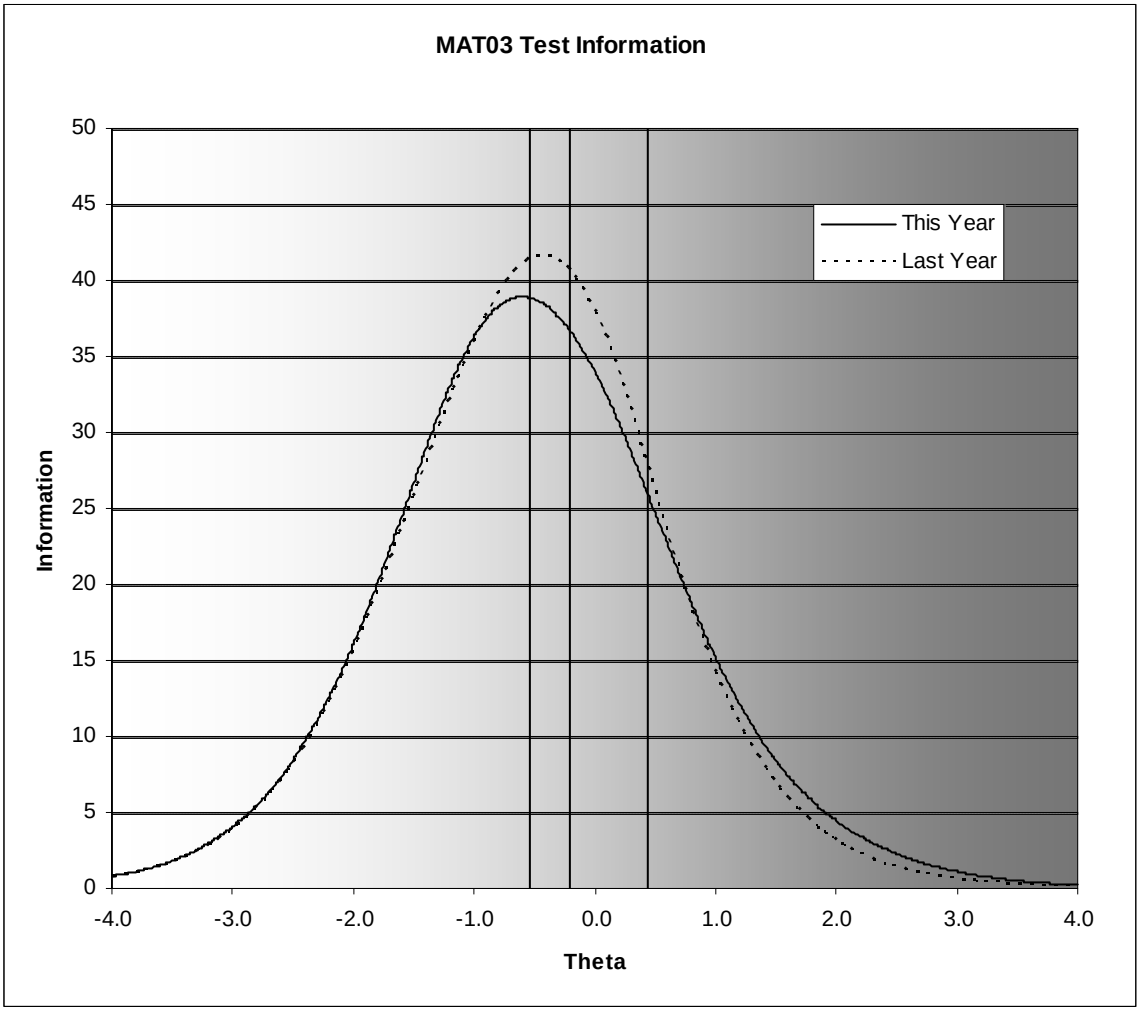


Figure H-3. 2009–10 MontCAS: TCC – Mathematics Grade 4

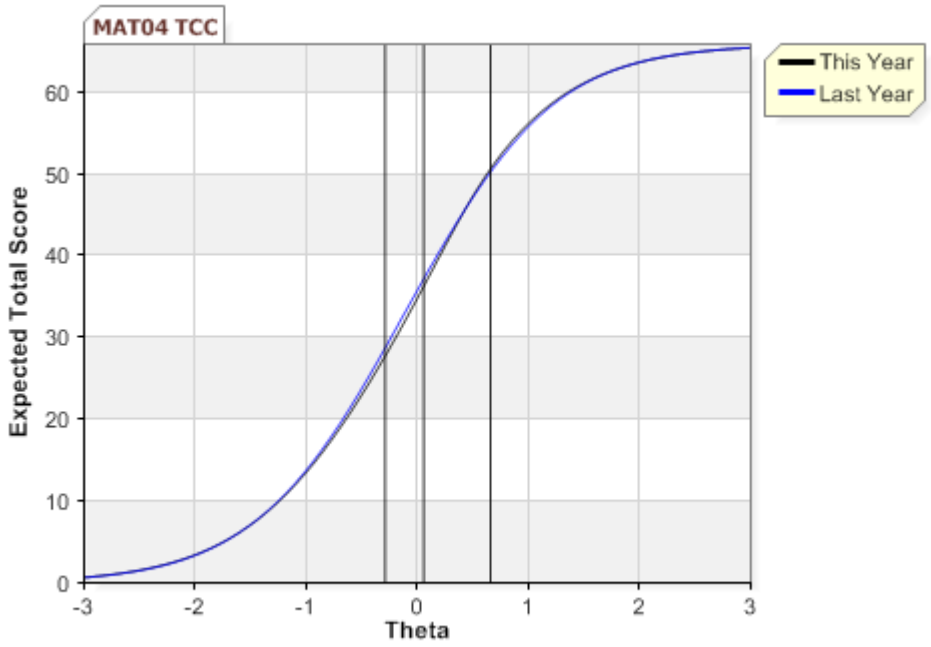


Figure H-4. 2009–10 MontCAS: TIF – Mathematics Grade 4

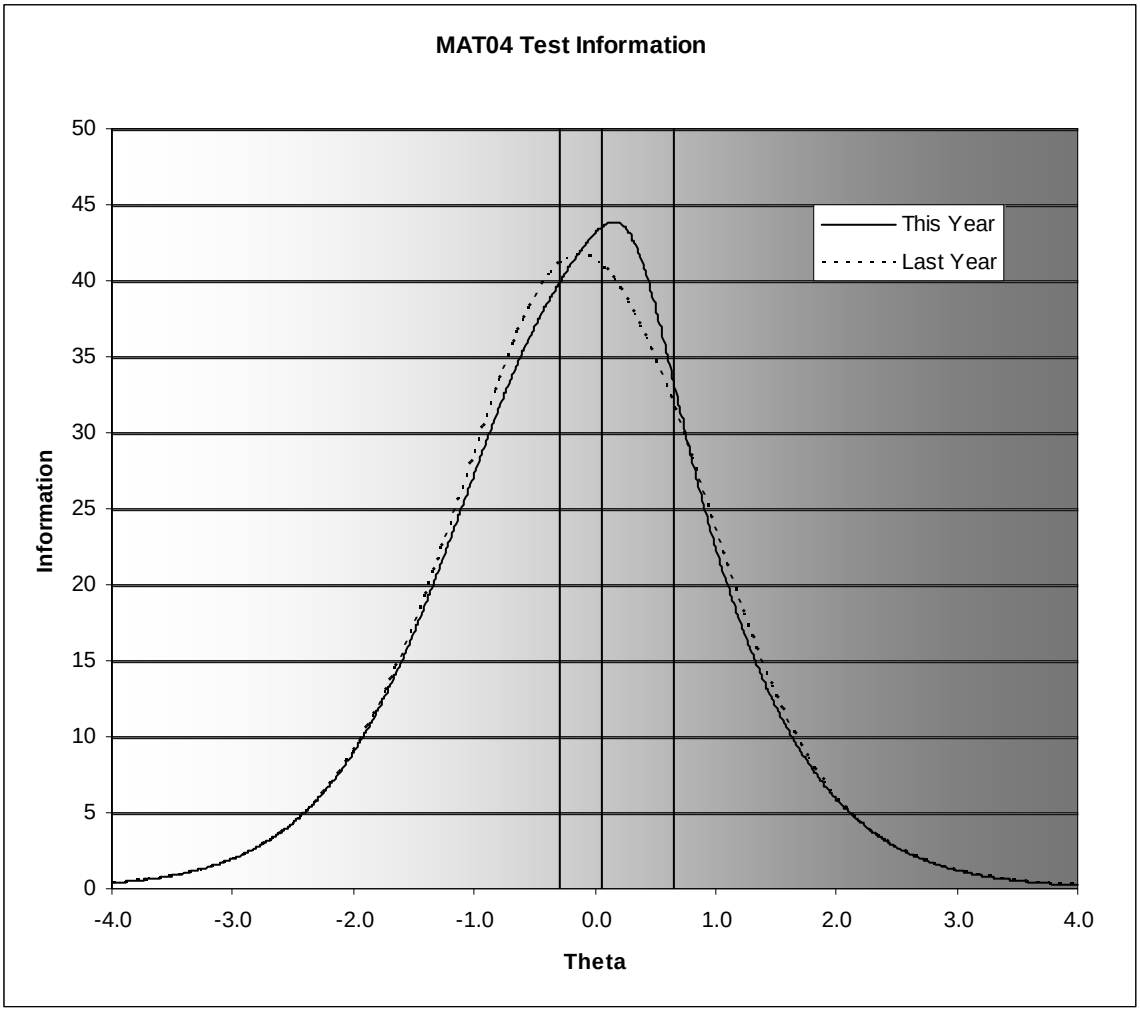


Figure H-5. 2009–10 MontCAS: TCC – Mathematics Grade 5

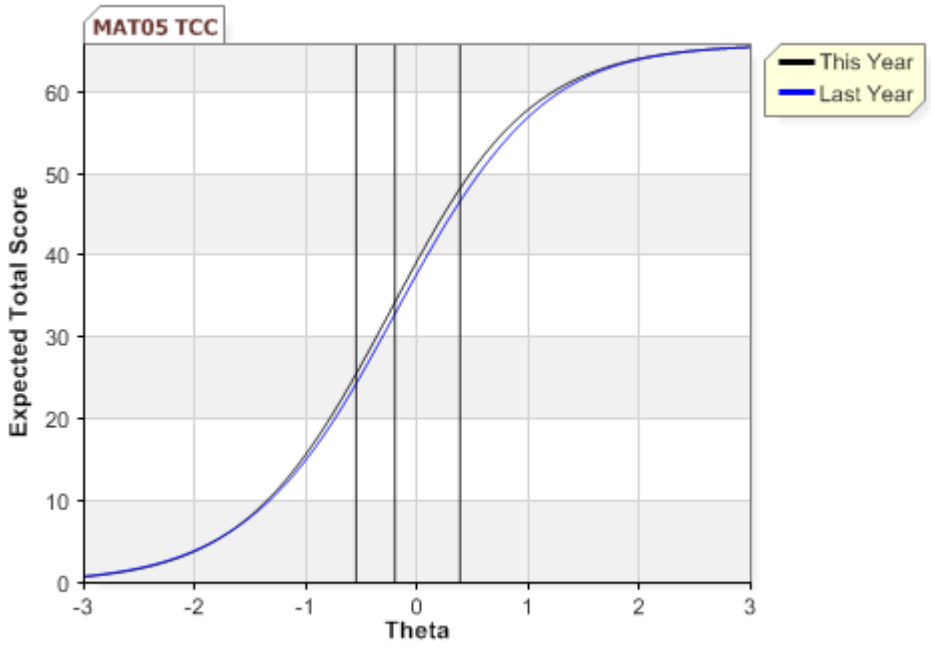


Figure H-6. 2009–10 MontCAS: TIF – Mathematics Grade 5

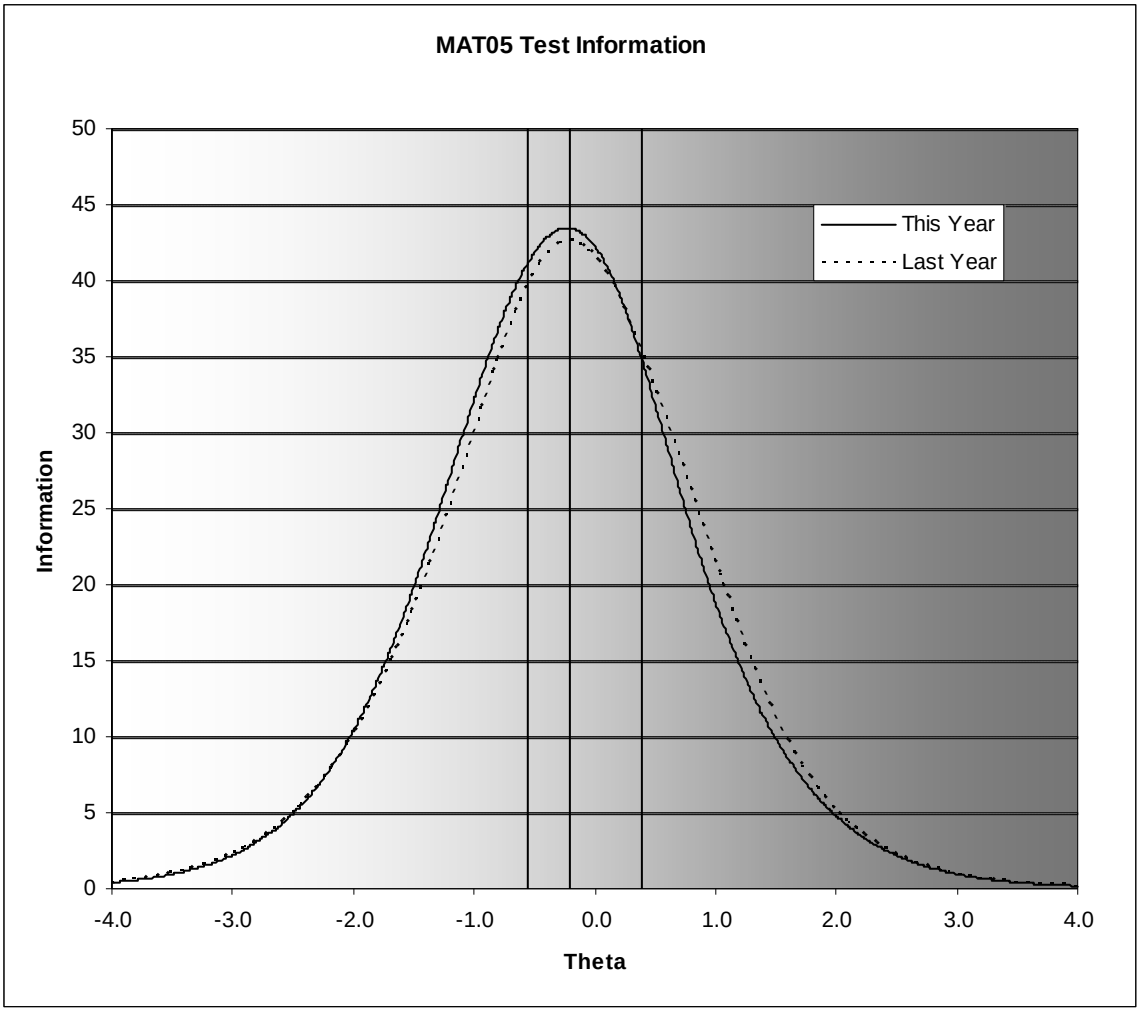


Figure H-7. 2009–10 MontCAS: TCC – Mathematics Grade 6

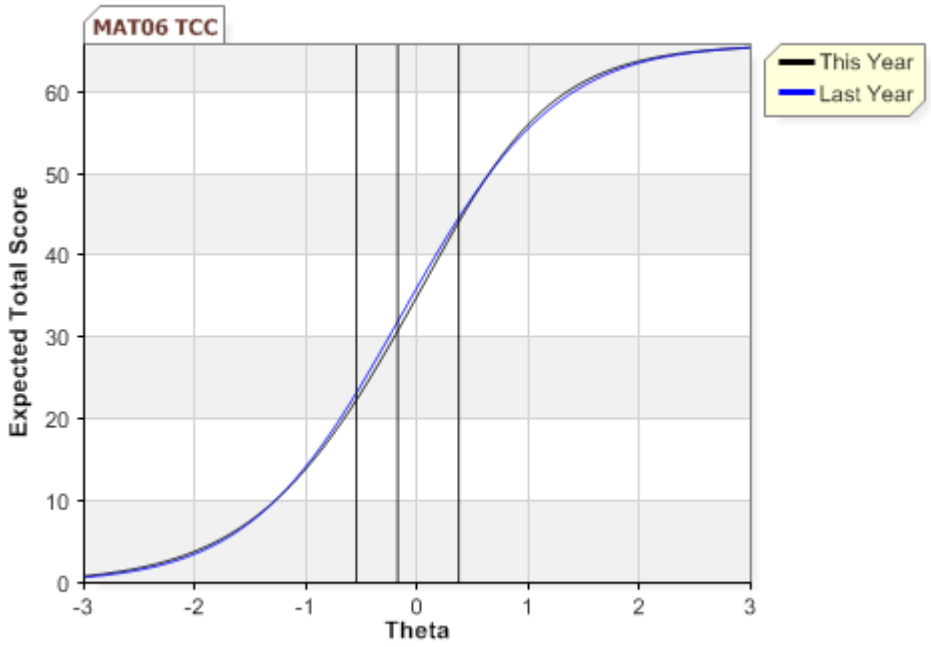


Figure H-8. 2009–10 MontCAS: TIF – Mathematics Grade 6

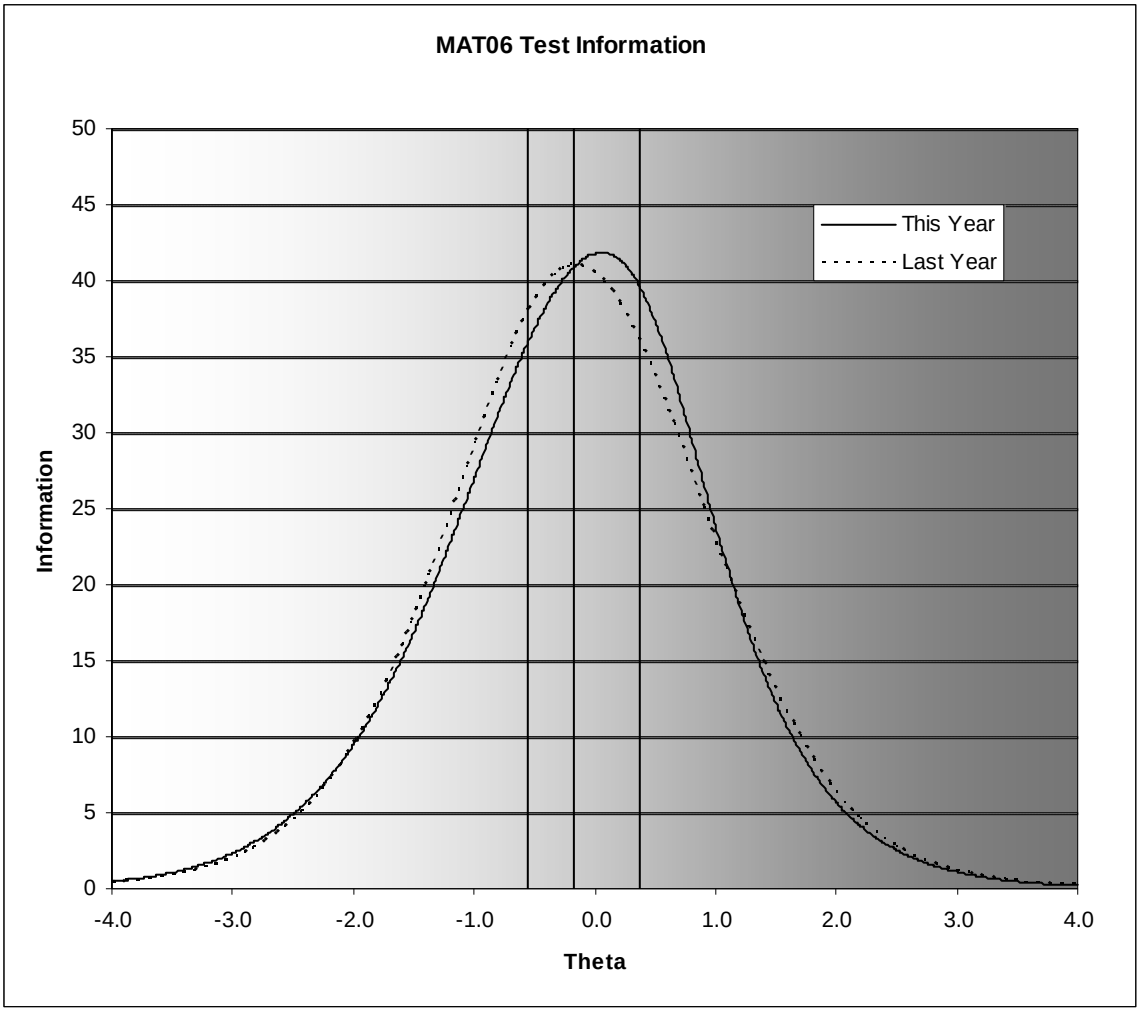


Figure H-9. 2009–10 MontCAS: TCC – Mathematics Grade 7

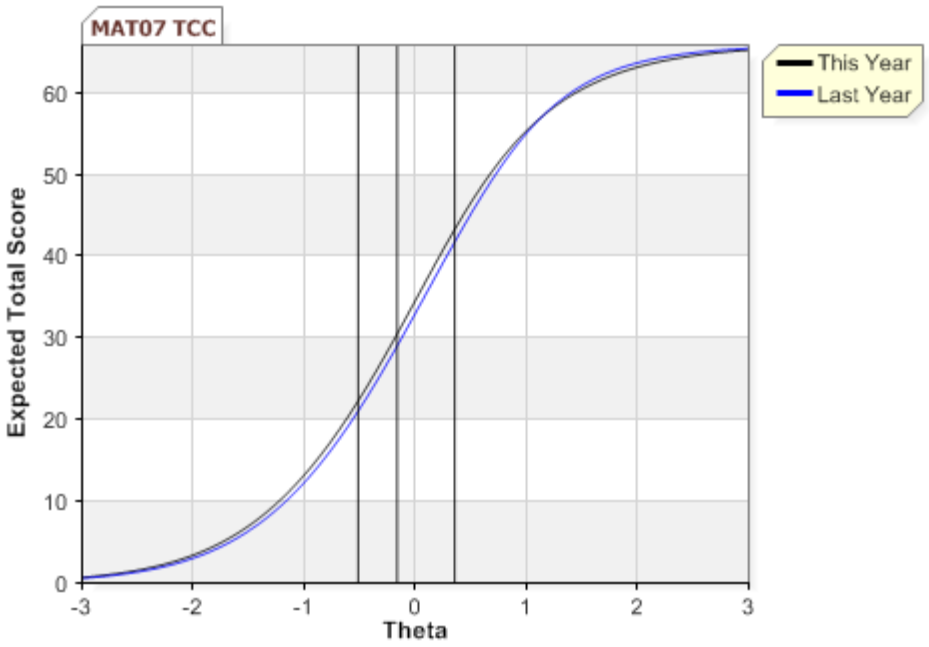


Figure H-10. 2009–10 MontCAS: TIF – Mathematics Grade 7

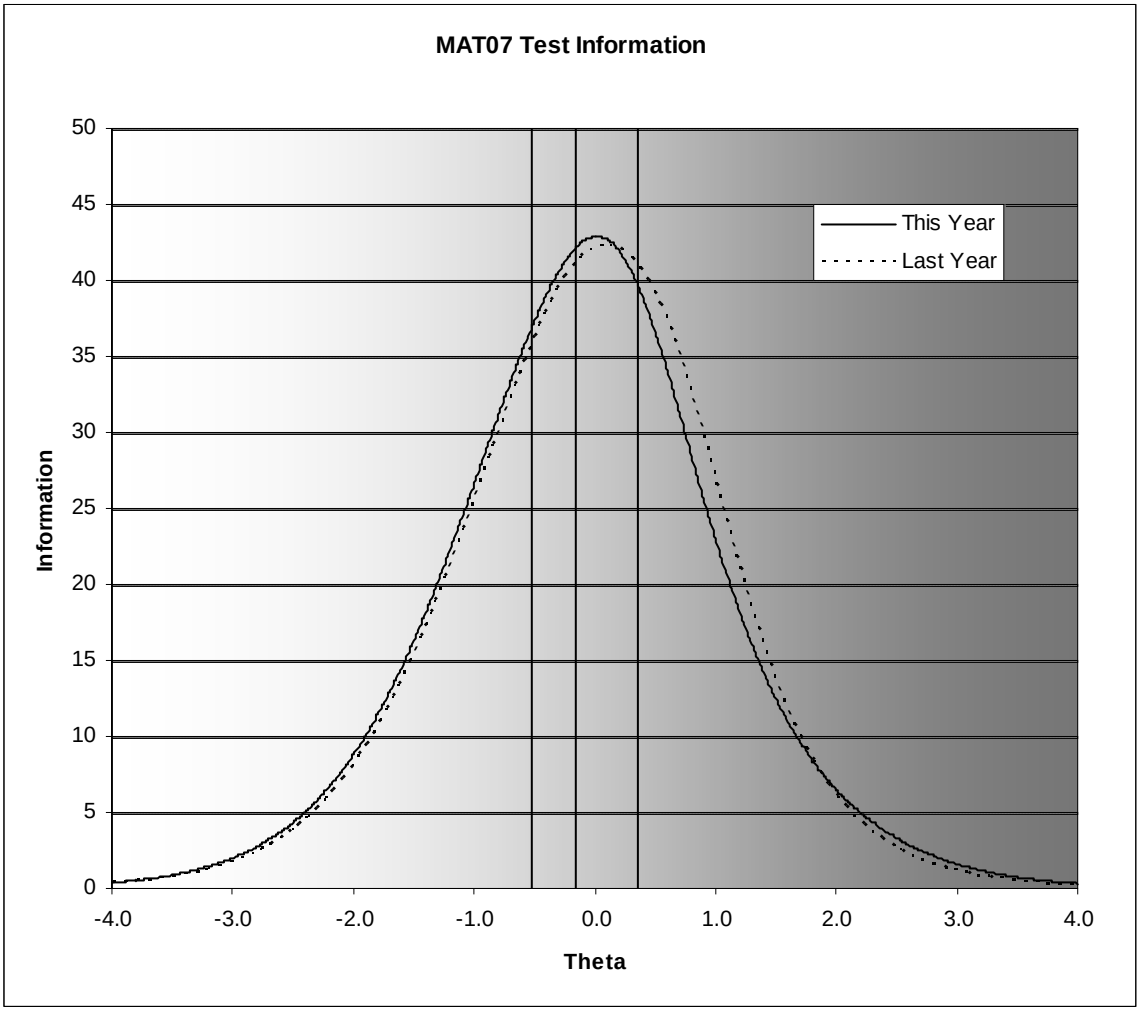


Figure H-11. 2009–10 MontCAS: TCC – Mathematics Grade 8

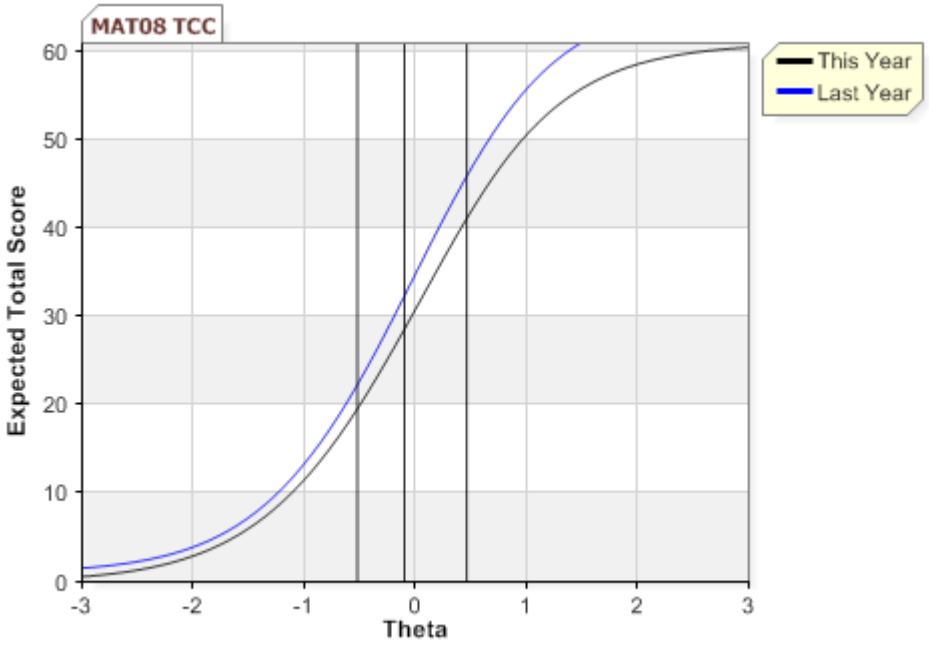


Figure H-12. 2009–10 MontCAS: TIF – Mathematics Grade 8

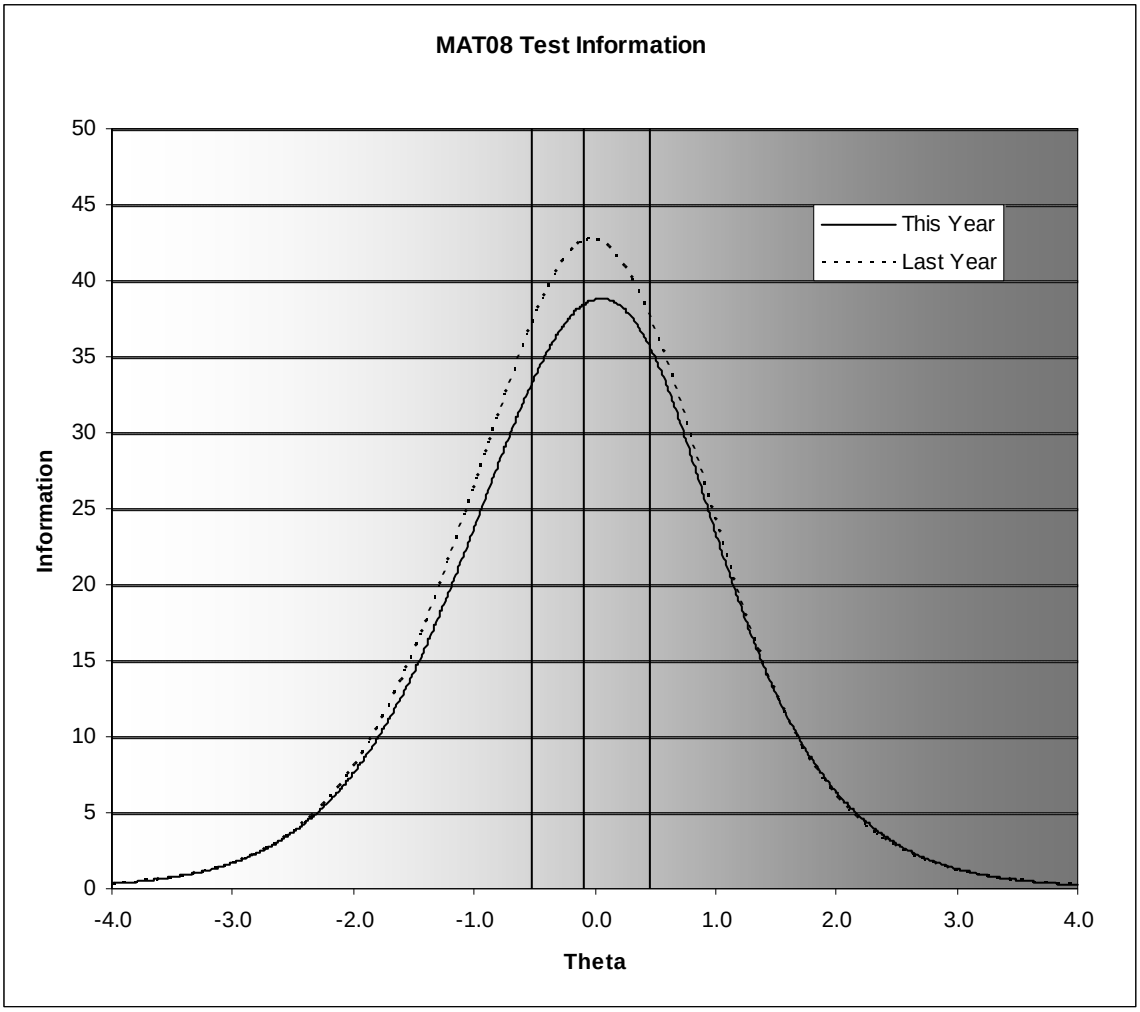


Figure H-13. 2009–10 MontCAS: TCC – Mathematics Grade 10

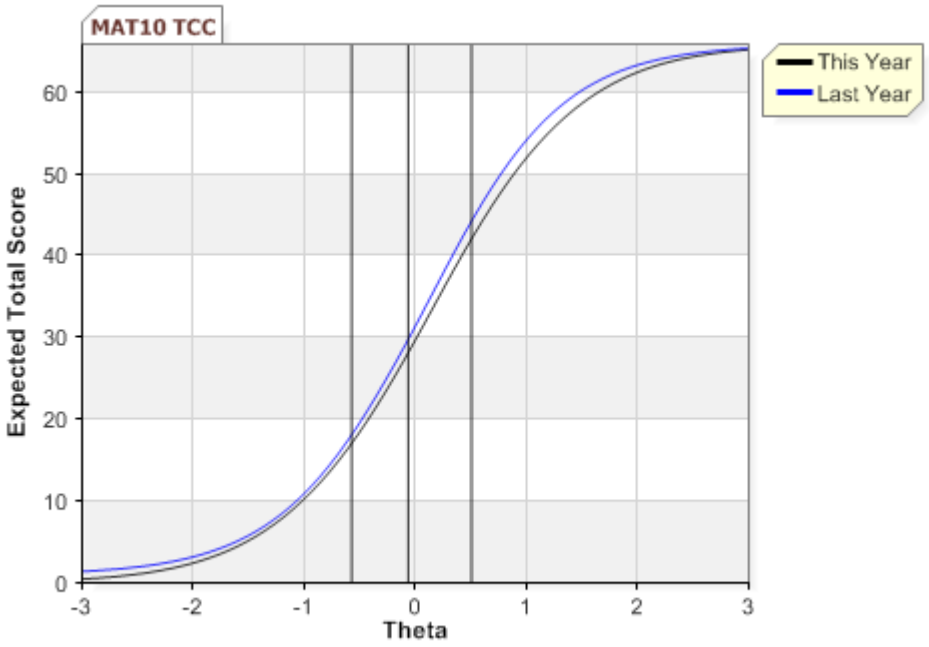


Figure H-14. 2009–10 MontCAS: TIF – Mathematics Grade 10

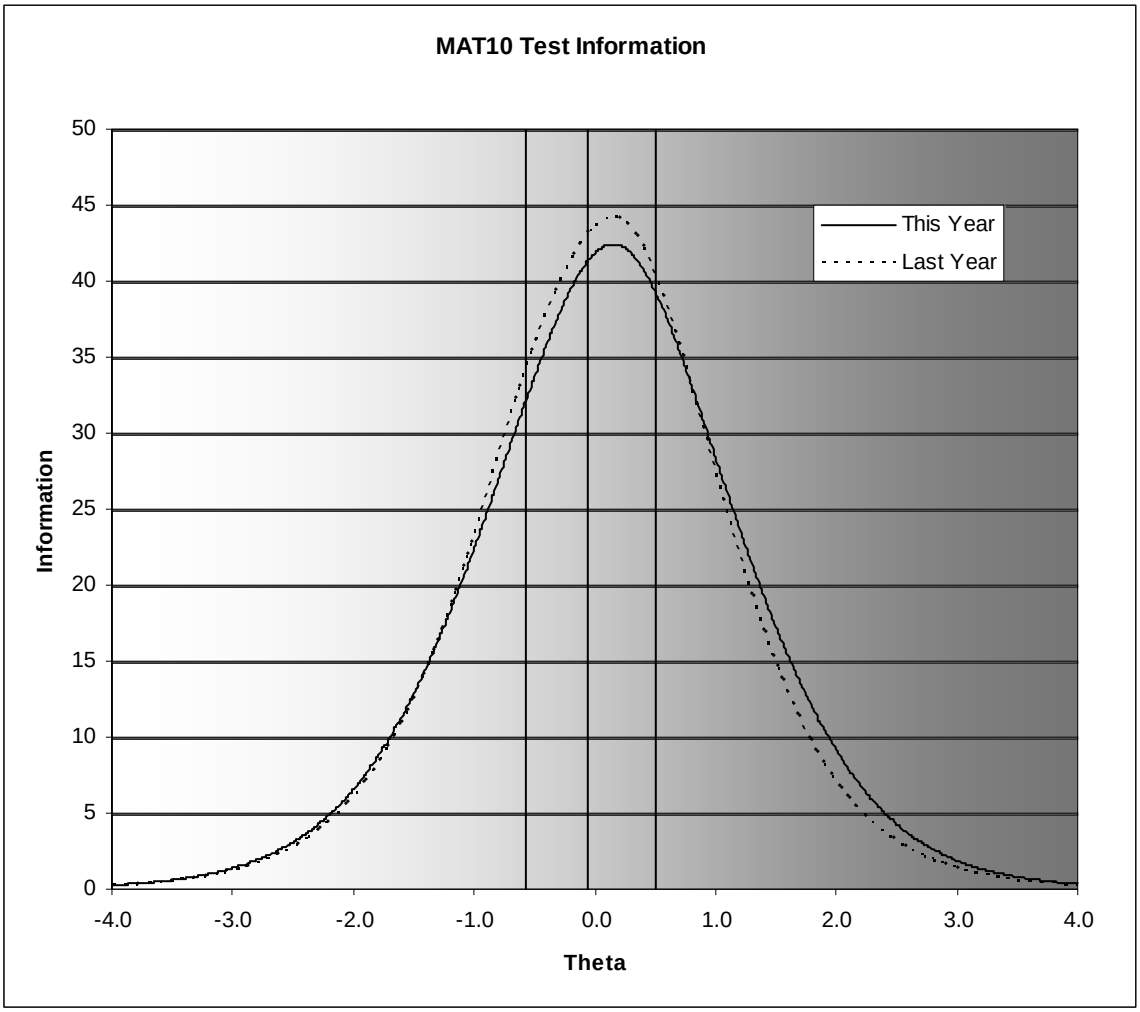


Figure H-15. 2009–10 MontCAS: TCC – Reading Grade 3

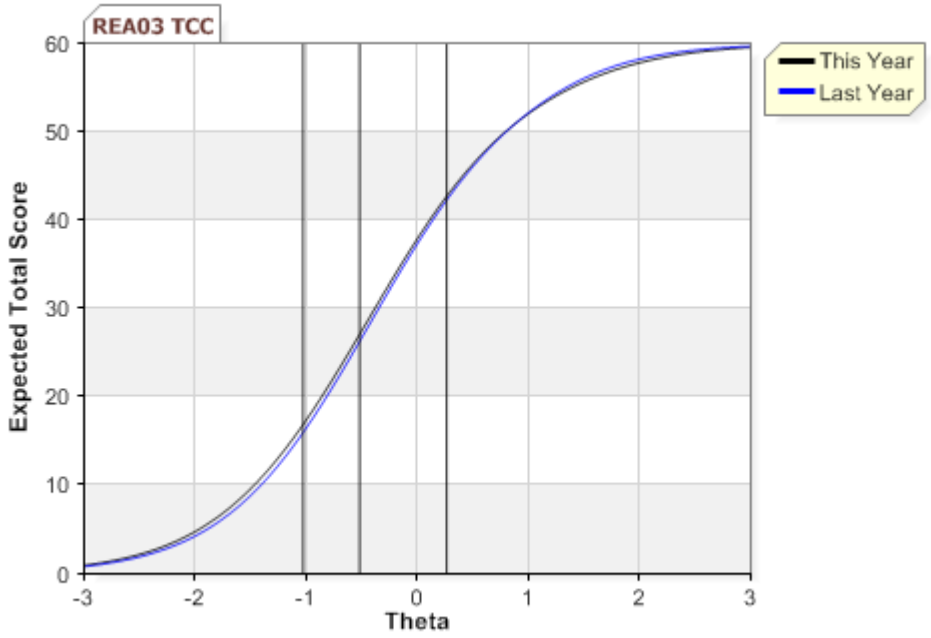


Figure H-16. 2009–10 MontCAS: TIF – Reading Grade 3

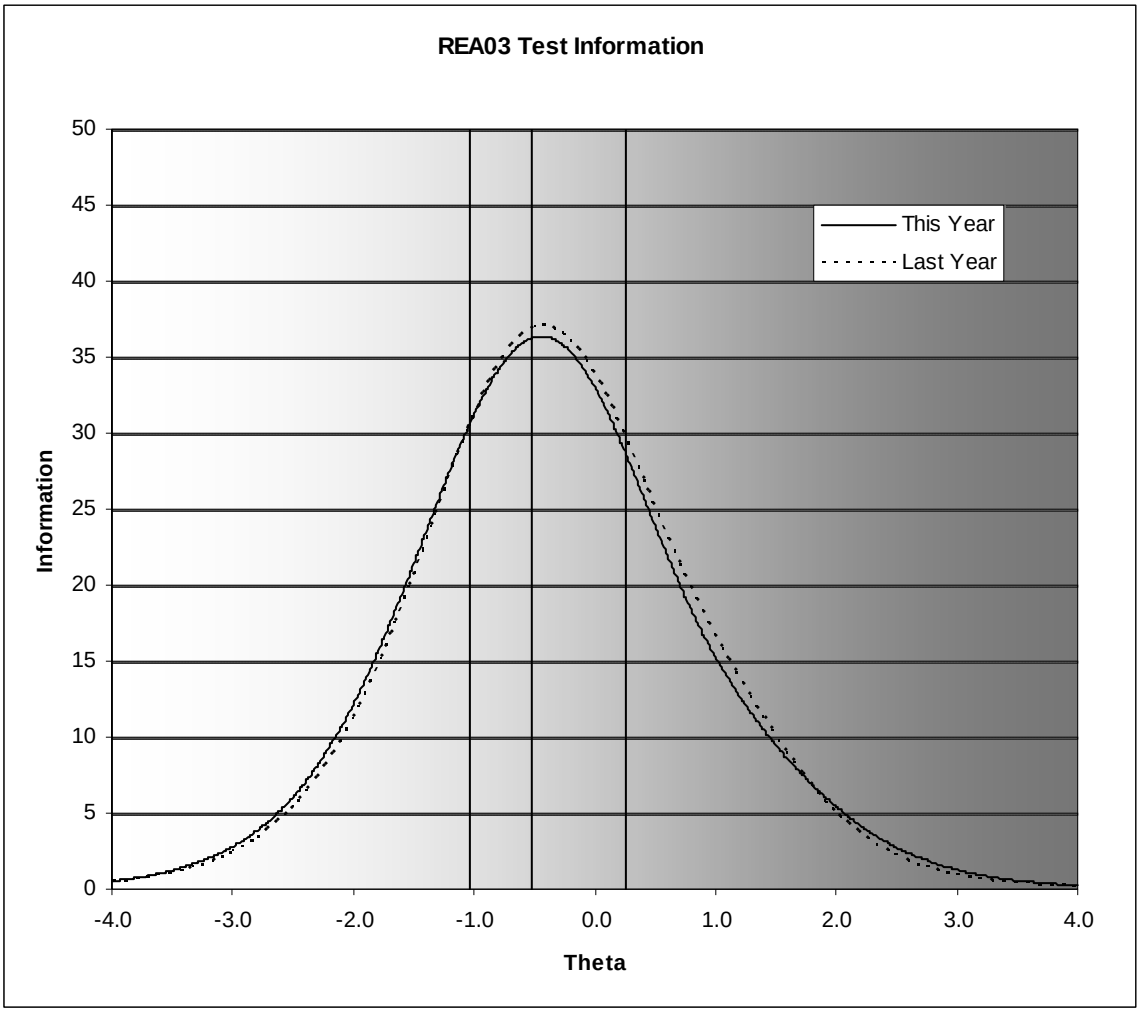


Figure H-17. 2009–10 MontCAS: TCC – Reading Grade 4

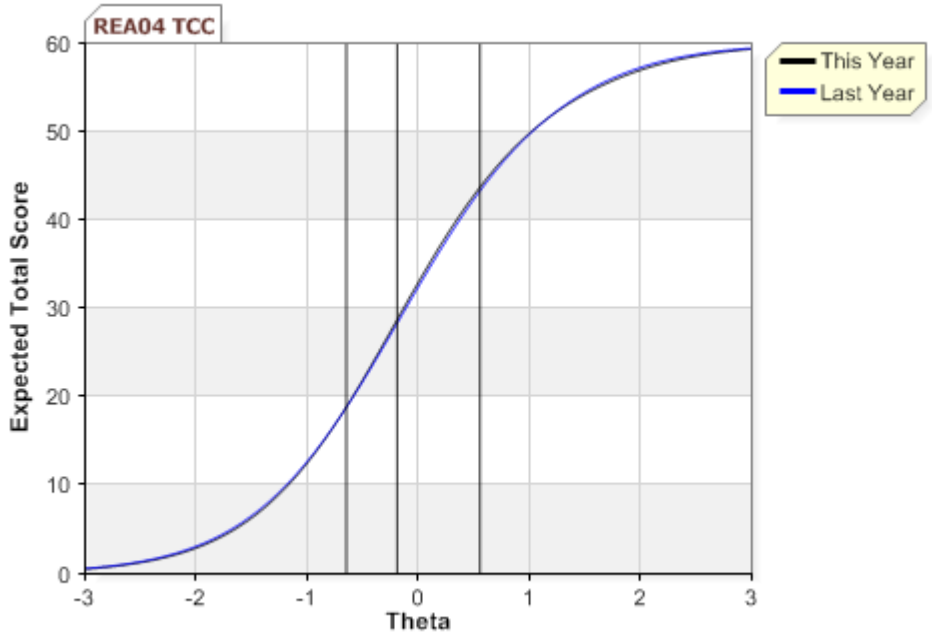


Figure H-18. 2009–10 MontCAS: TIF – Reading Grade 4

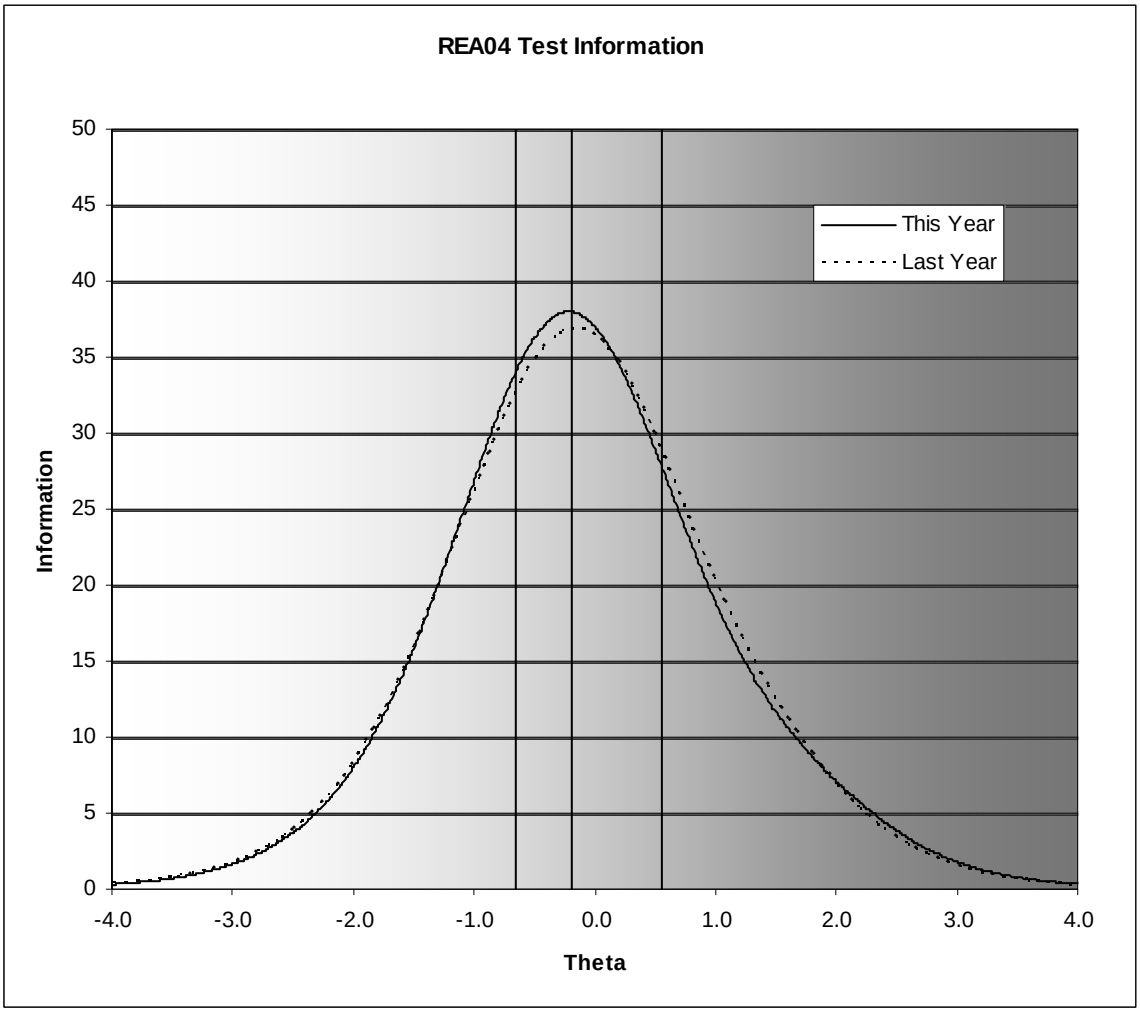


Figure H-19. 2009–10 MontCAS: TCC – Reading Grade 5

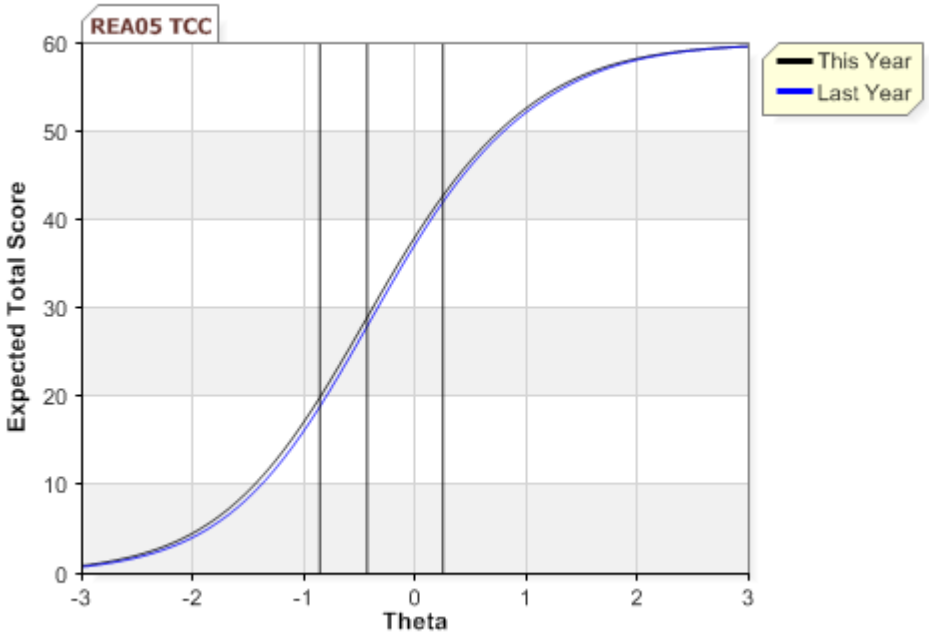


Figure H-20. 2009–10 MontCAS: TIF – Reading Grade 5

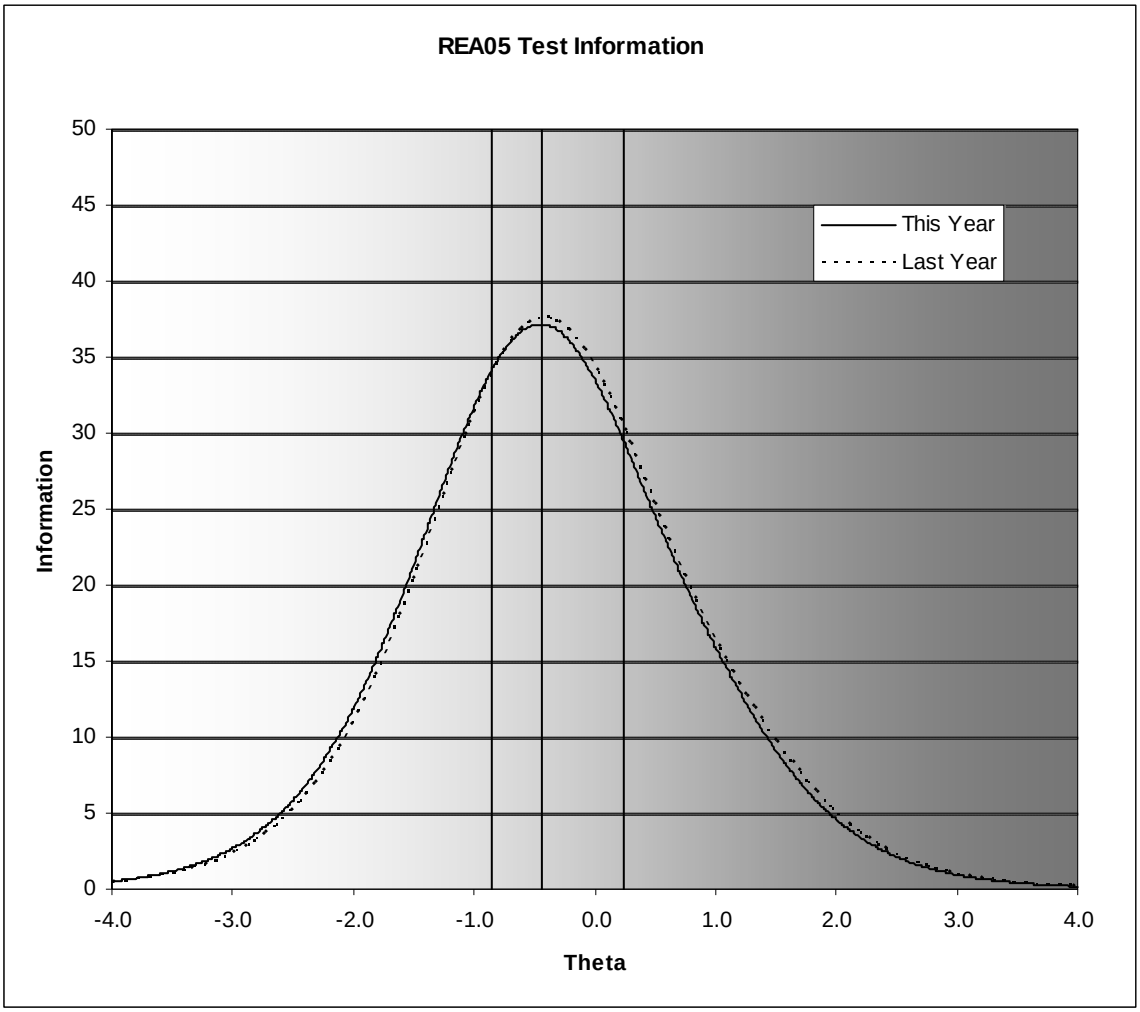


Figure H-21. 2009–10 MontCAS: TCC – Reading Grade 6

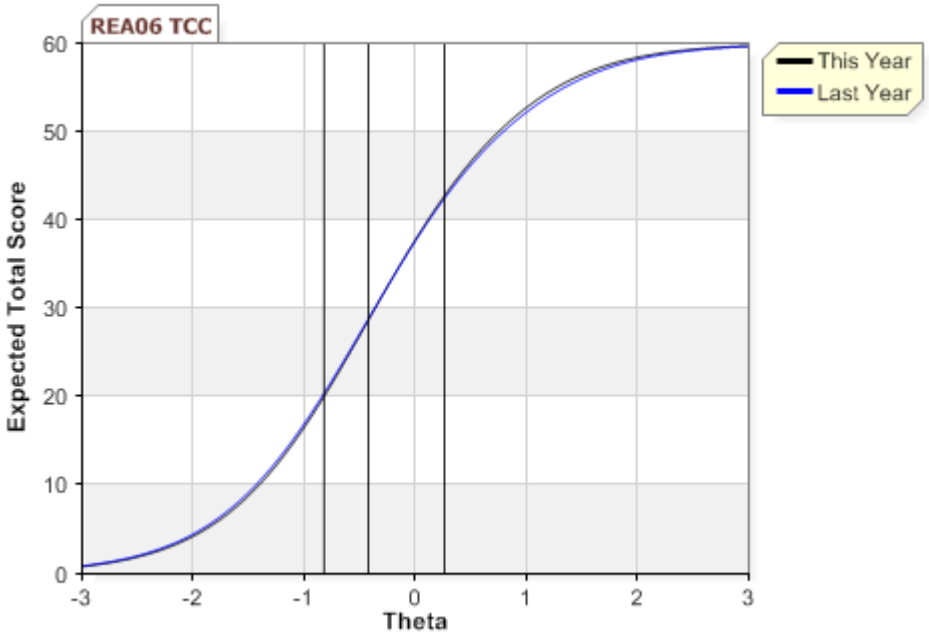


Figure H-22. 2009–10 MontCAS: TIF – Reading Grade 6

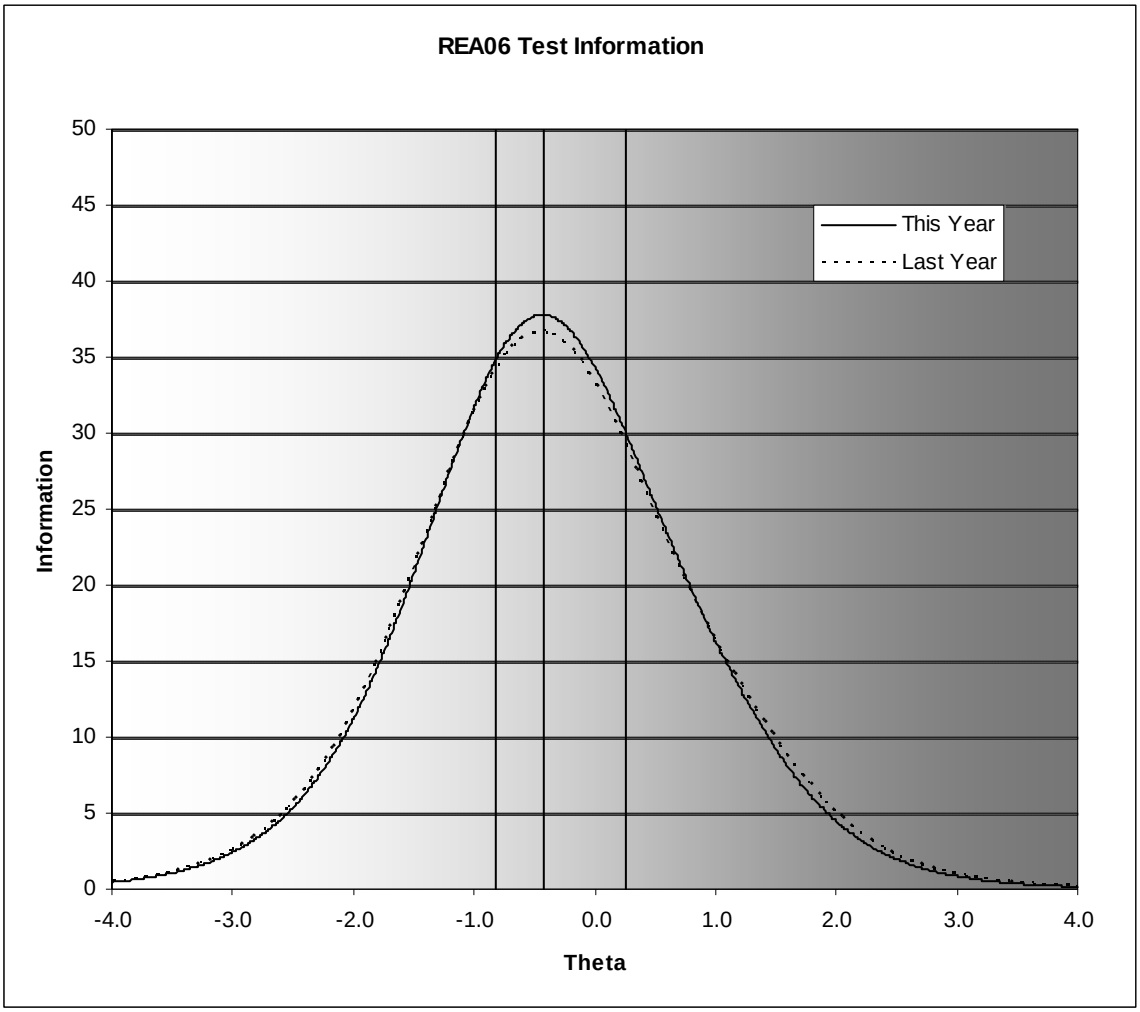


Figure H-23. 2009–10 MontCAS: TCC – Reading Grade 7

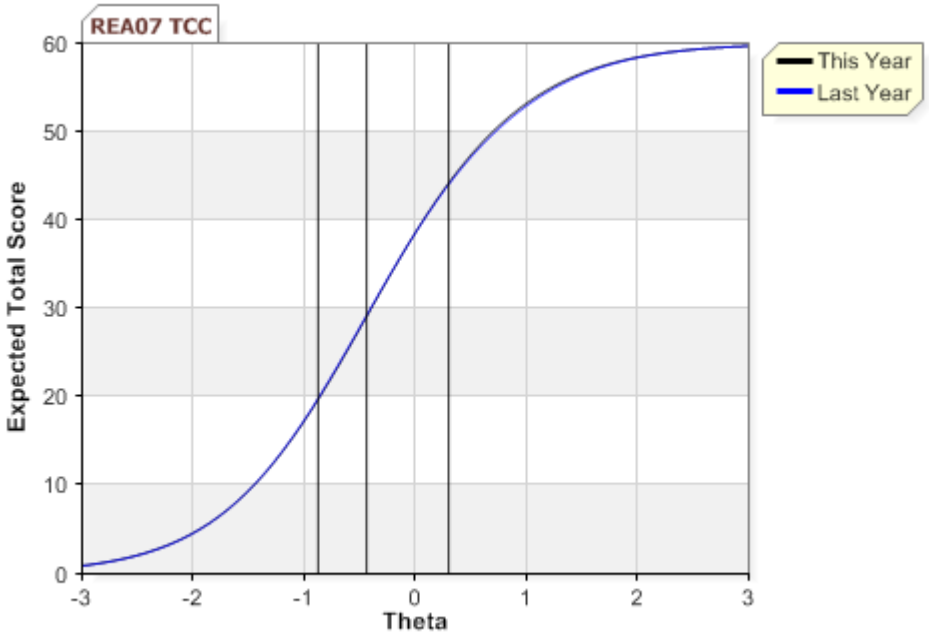


Figure H-24. 2009–10 MontCAS: TIF – Reading Grade 7

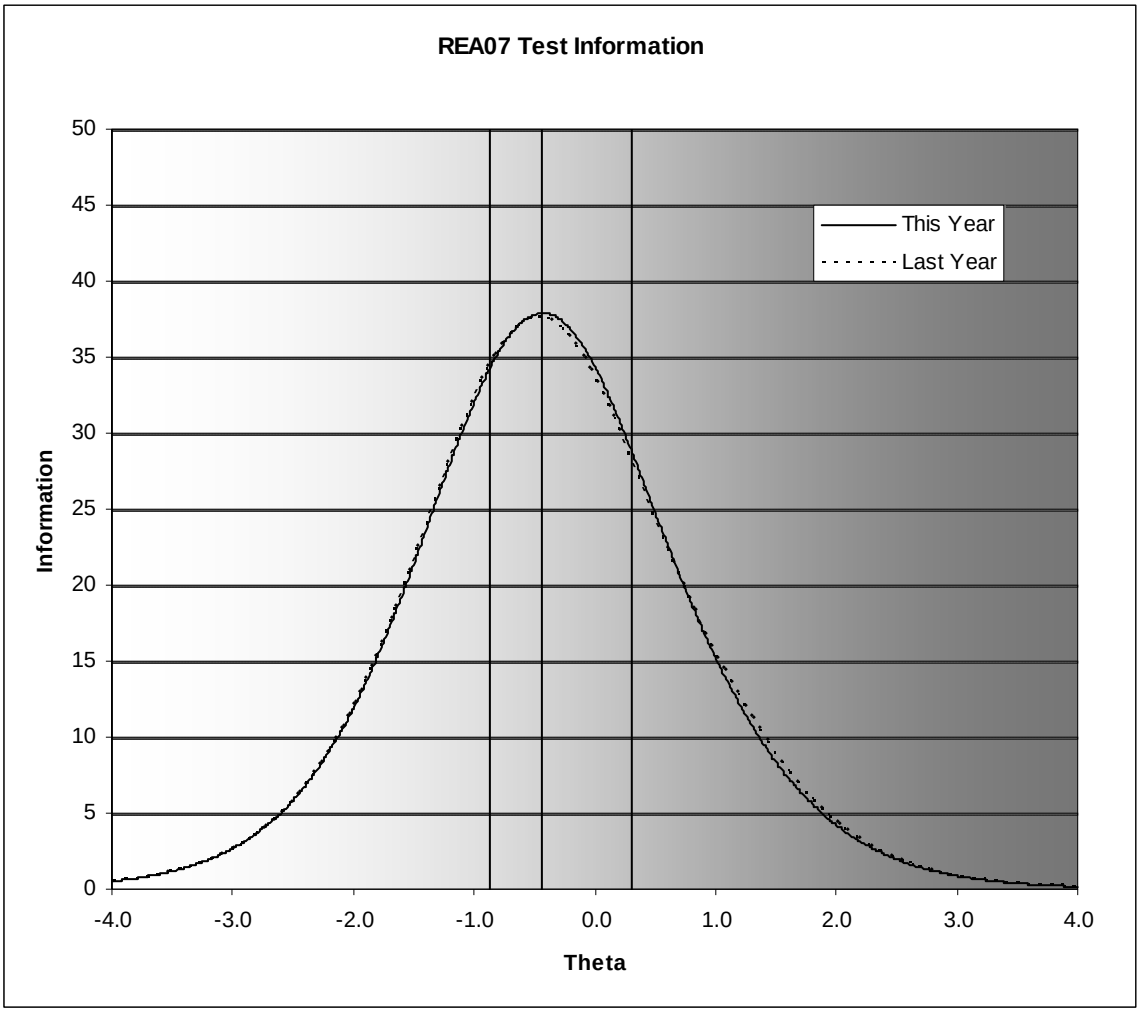


Figure H-25. 2009–10 MontCAS: TCC – Reading Grade 8

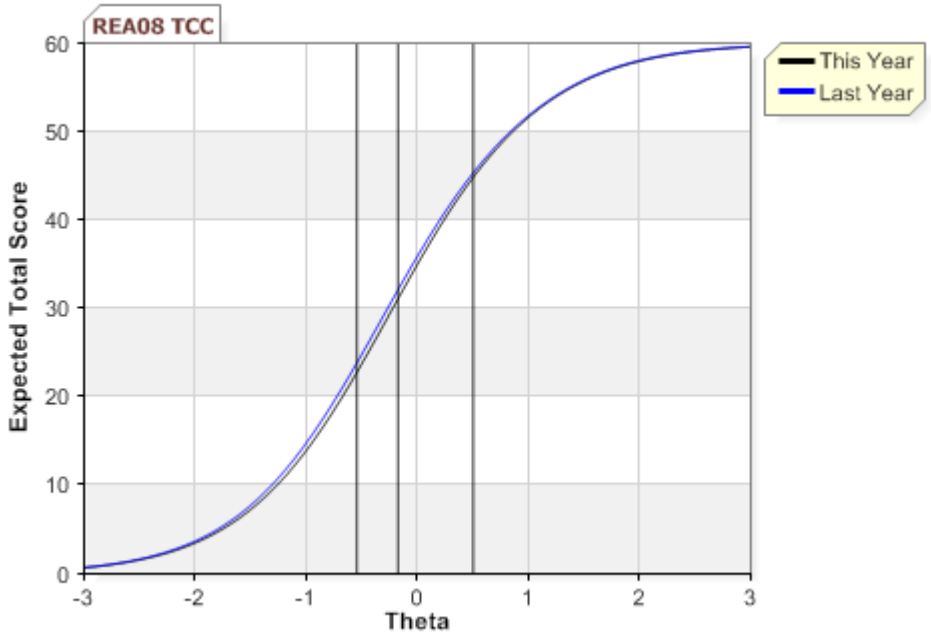


Figure H-26. 2009–10 MontCAS: TIF – Reading Grade 8

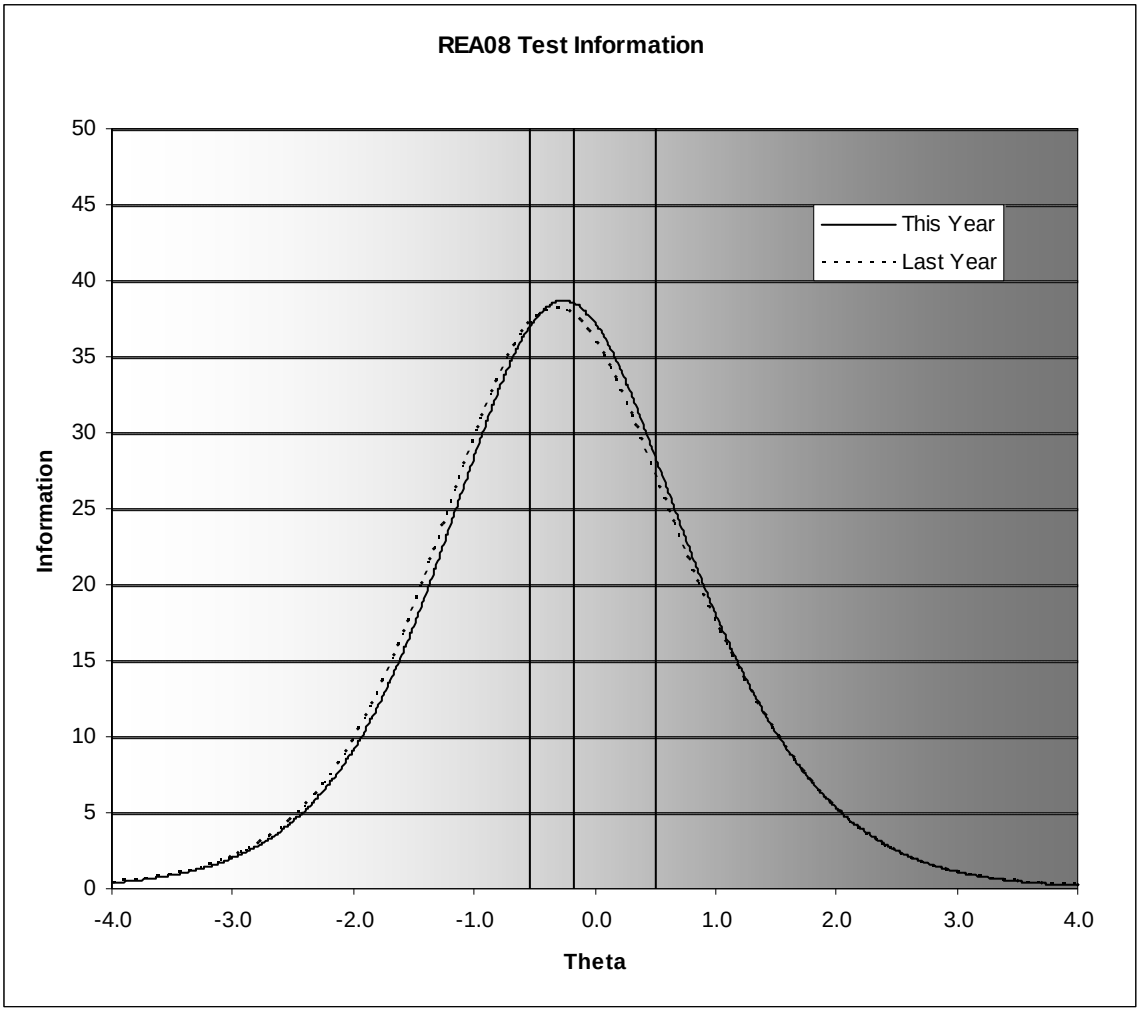


Figure H-27. 2009–10 MontCAS: TCC – Reading Grade 10

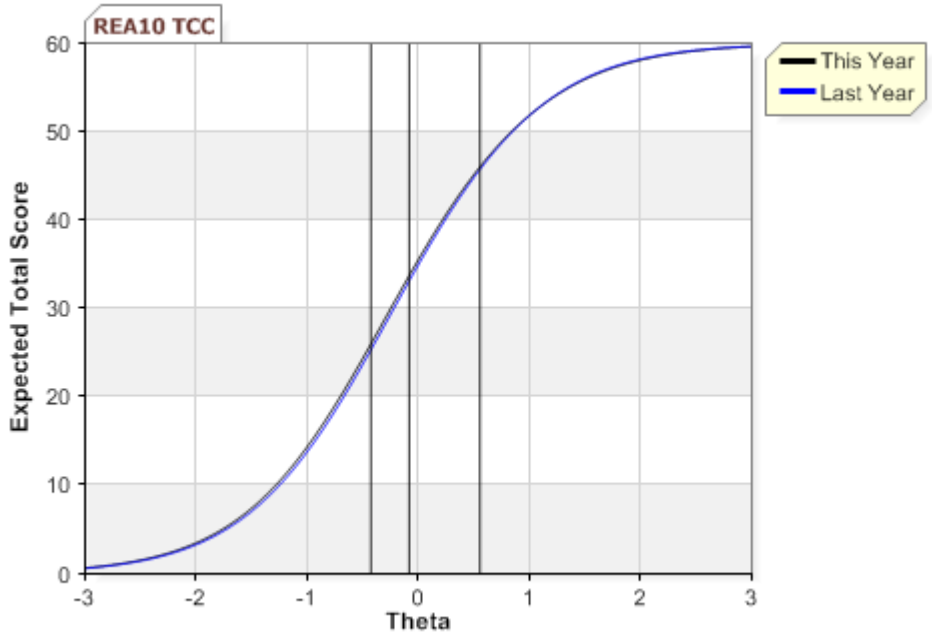


Figure H-28. 2009–10 MontCAS: TIF – Reading Grade 10

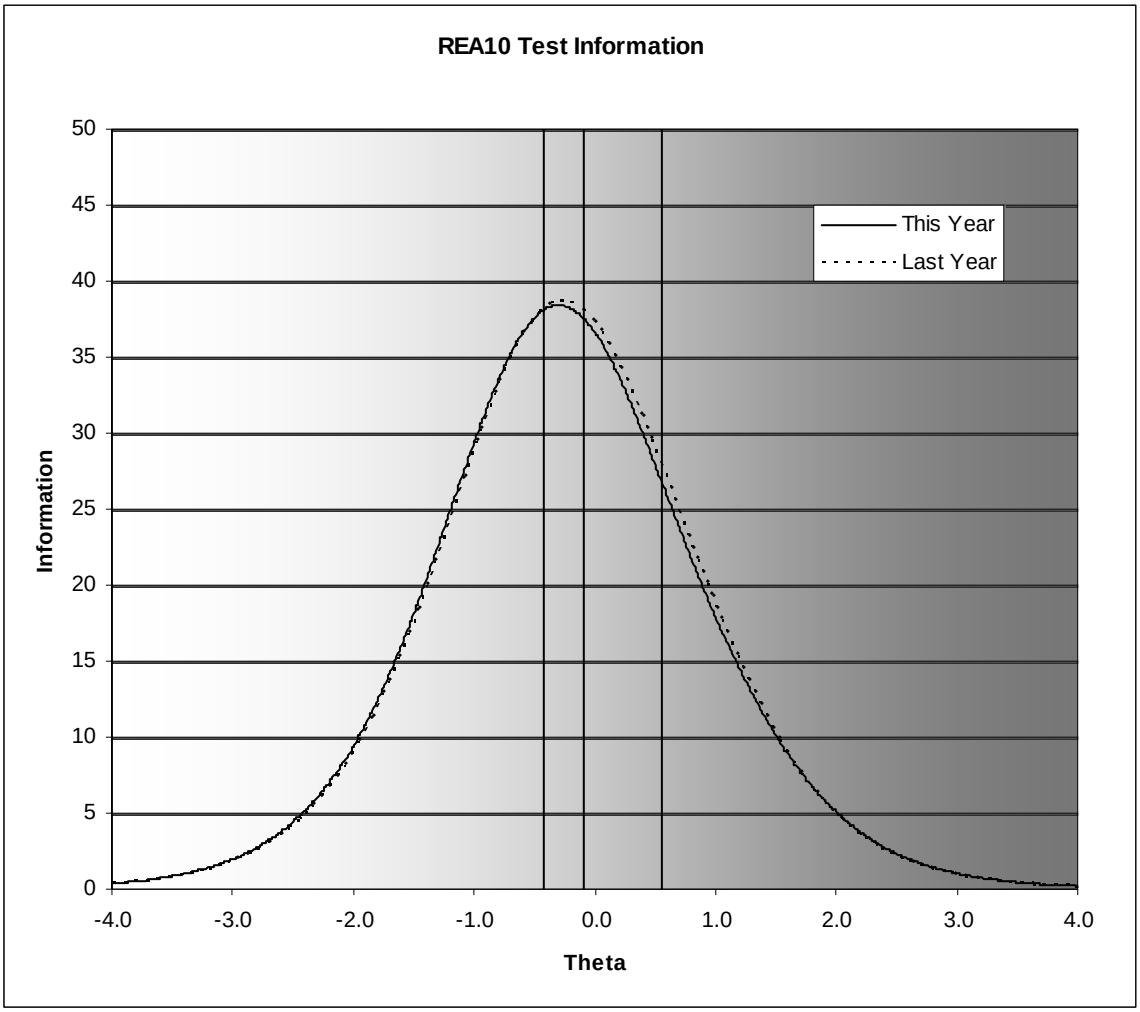


Figure H-29. 2009–10 MontCAS: TCC – Science Grade 4

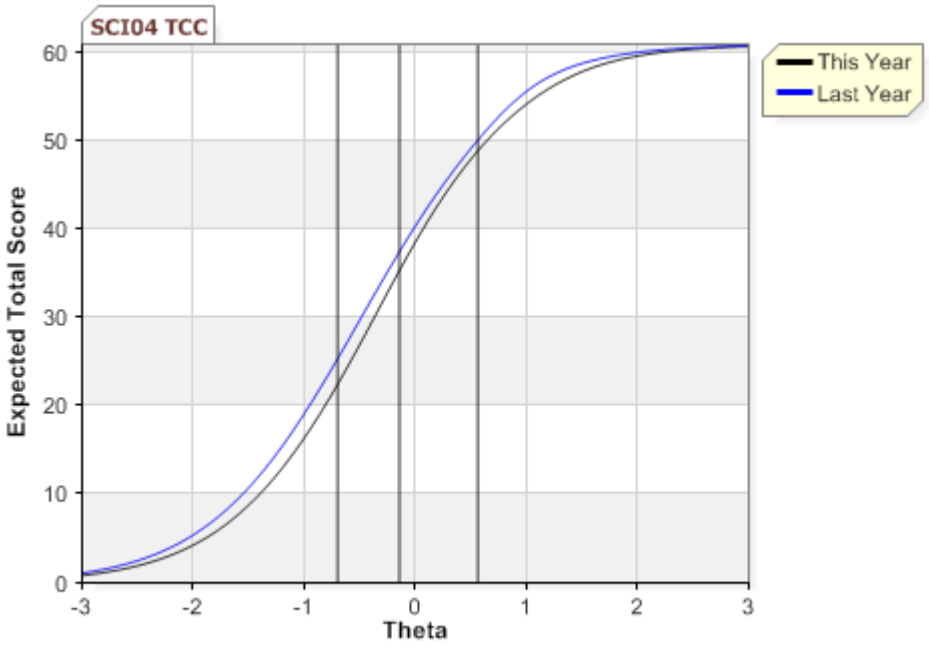


Figure H-30. 2009–10 MontCAS: TIF – Science Grade 4

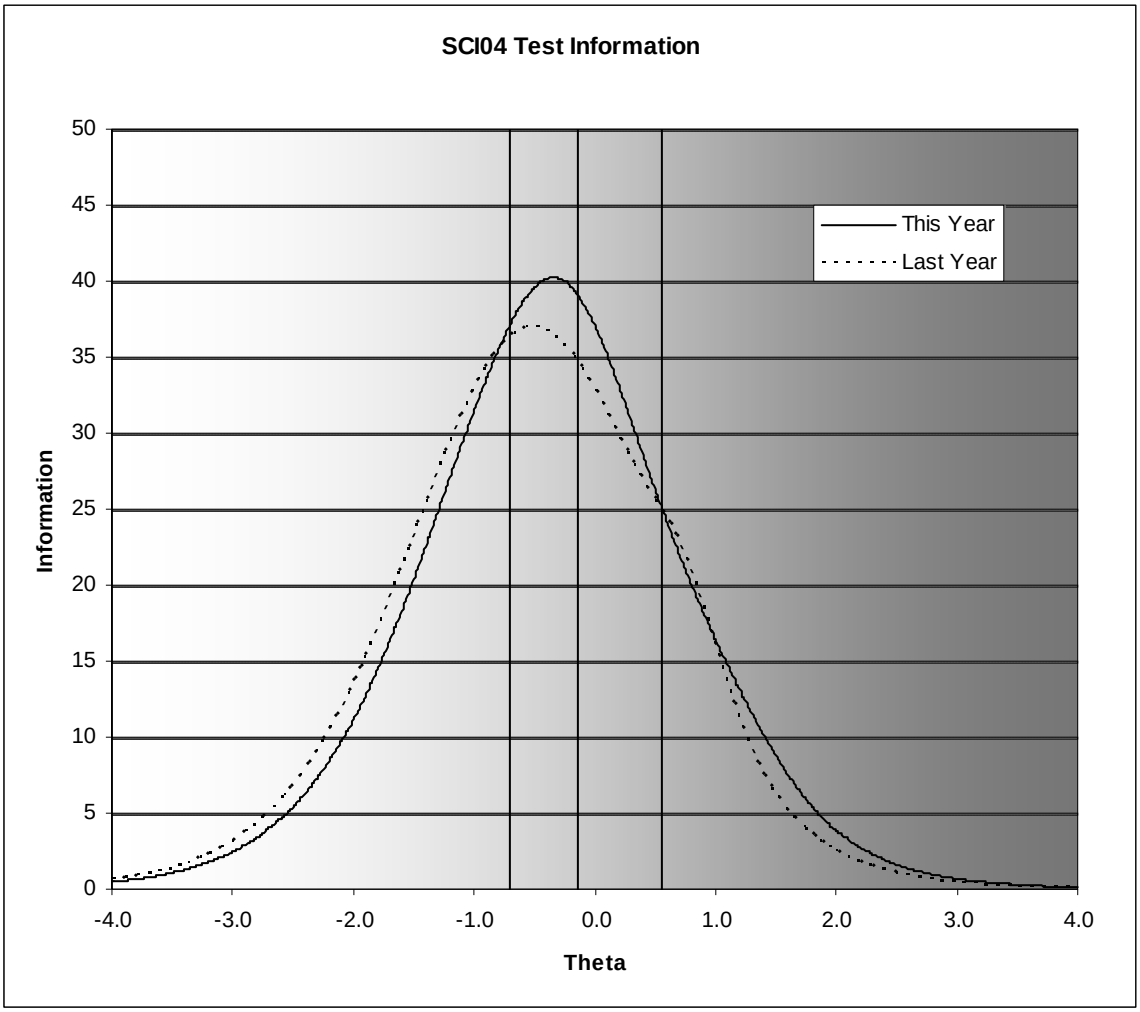


Figure H-31. 2009–10 MontCAS: TCC – Science Grade 8

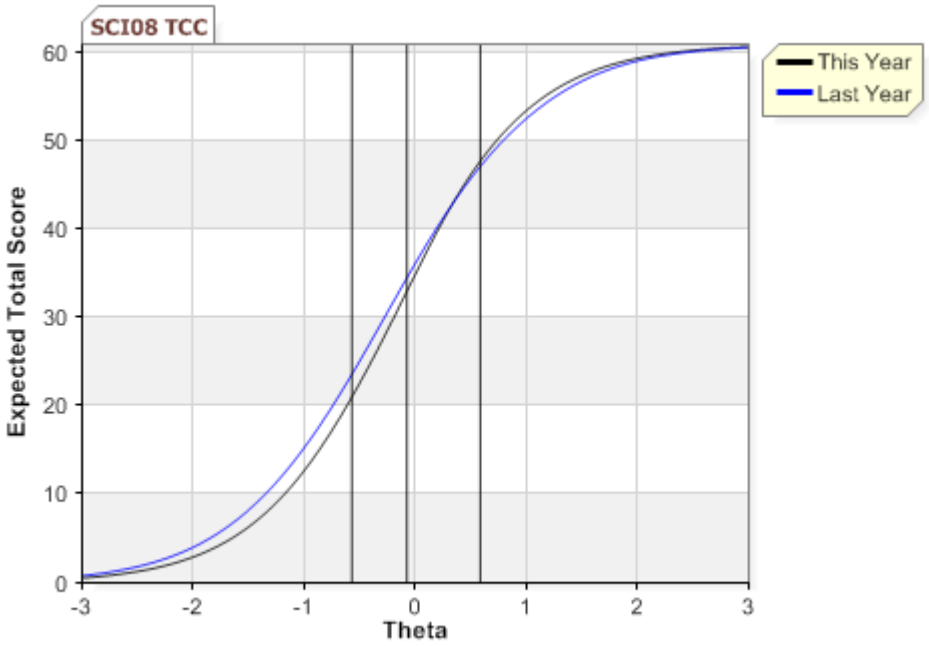


Figure H-32. 2009–10 MontCAS: TIF – Science Grade 8

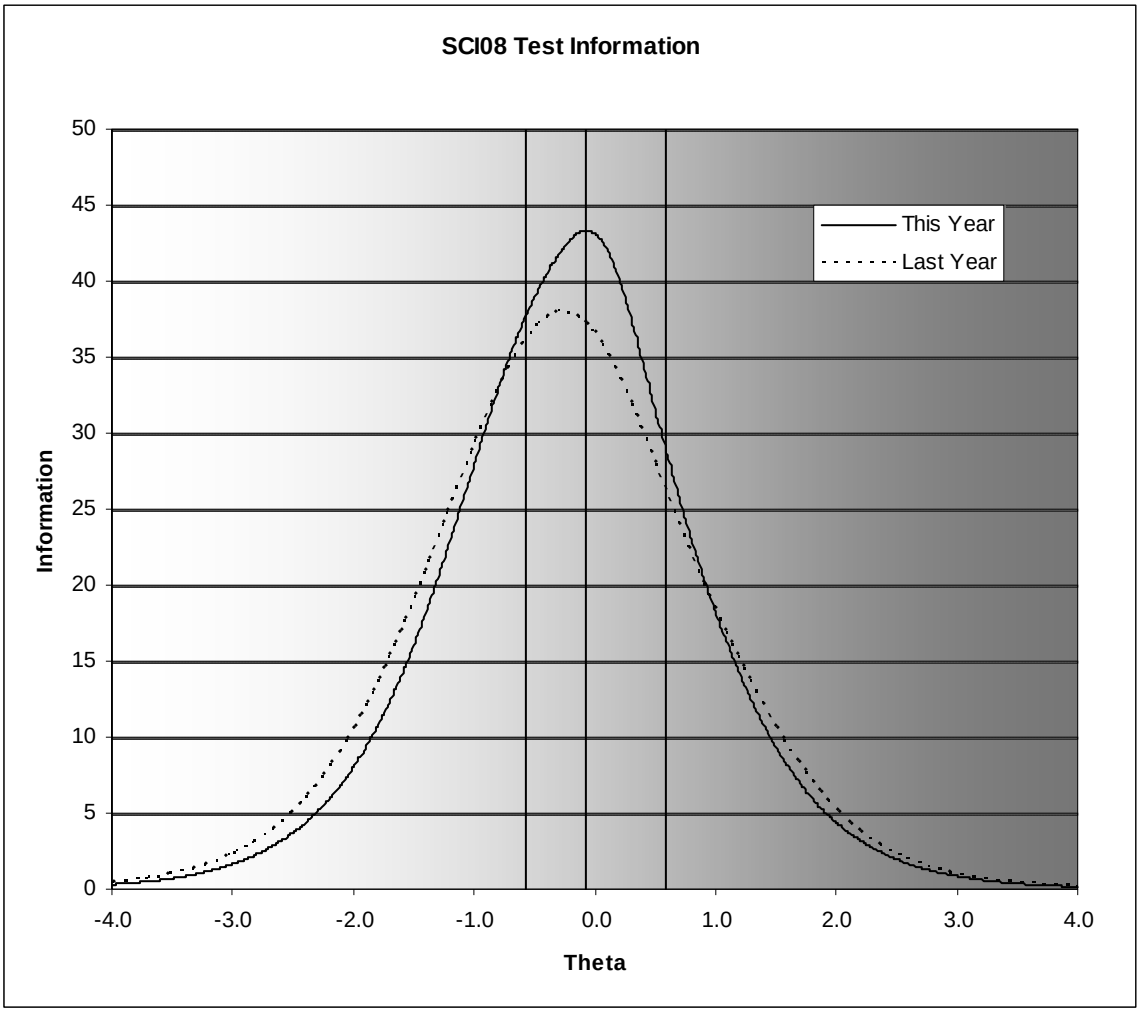


Figure H-33. 2009–10 MontCAS: TCC – Science Grade 10

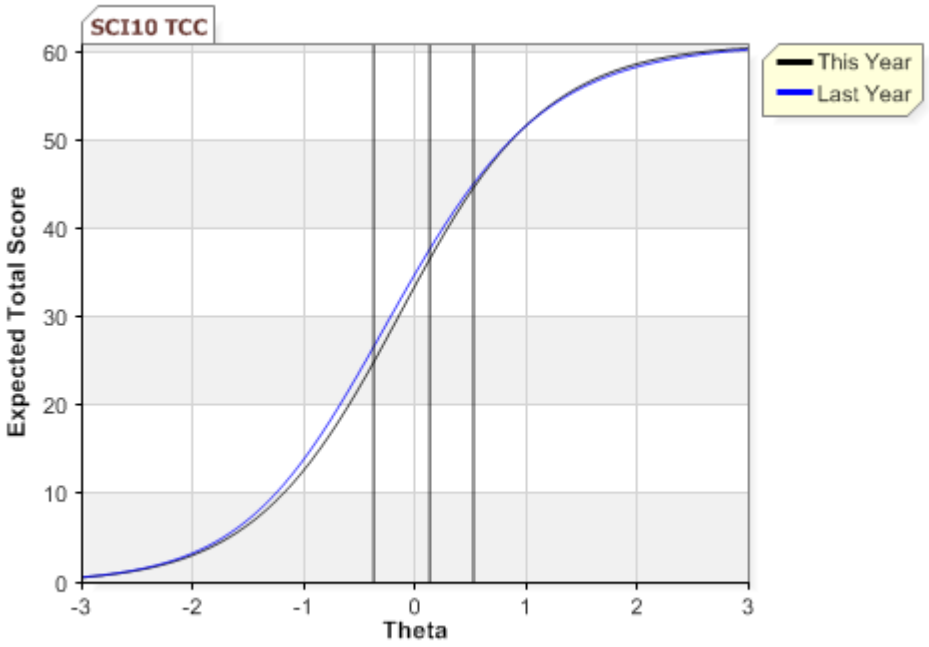
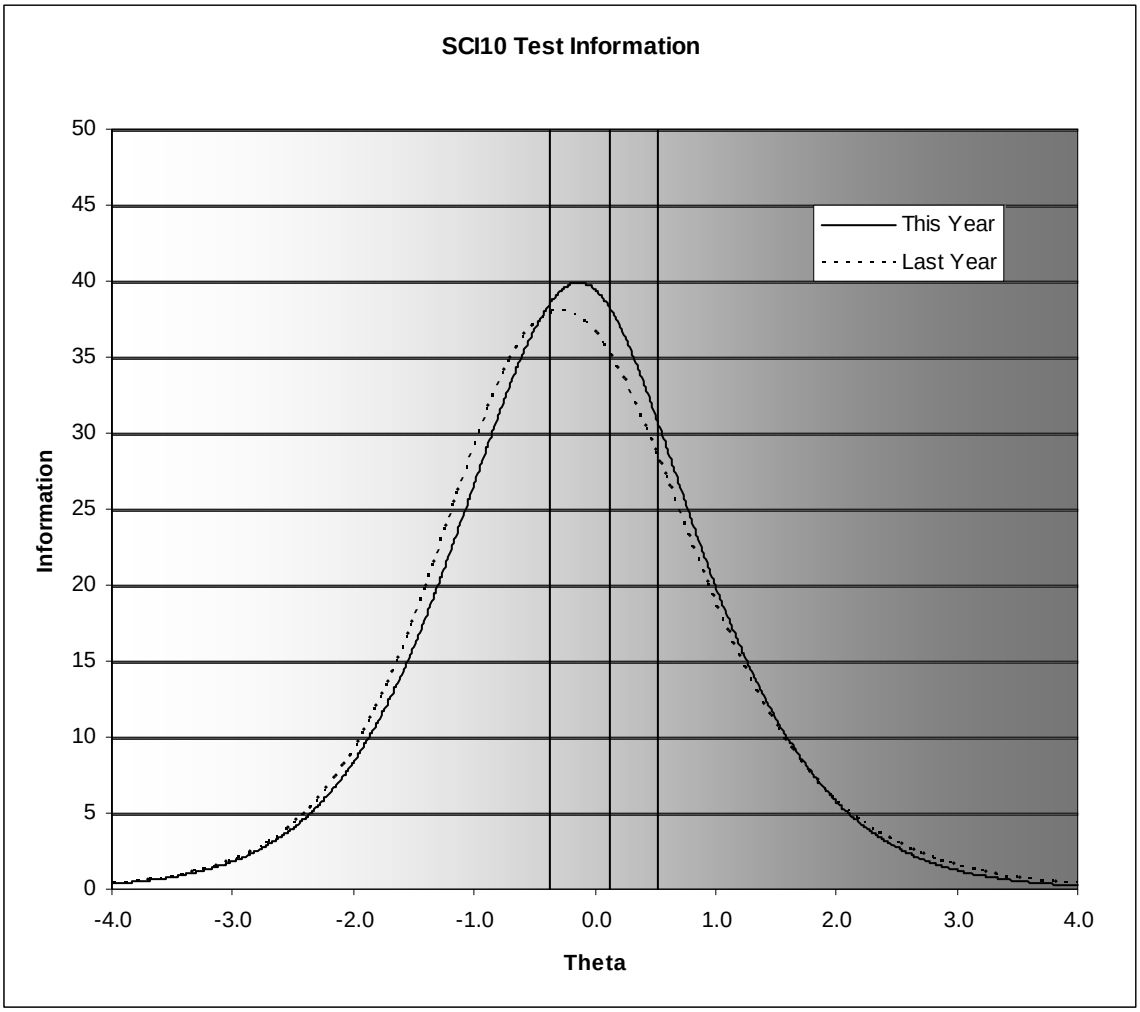


Figure H-34. 2009–10 MontCAS: TIF – Science Grade 10



Appendix I—Analyses of Equating Items (Delta & Rescore Analyses)

Table I-1. 2009–10 MontCAS: Delta Analysis Results – Mathematics Grade 3

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
42962	0.71	0.71	10.79	10.79	1	False	-0.31
42994	0.66	0.70	11.35	10.90	1	False	-0.15
43090	0.82	0.84	9.34	9.02	1	False	-1.15
43105	0.63	0.68	11.67	11.13	1	False	0.32
43136	0.70	0.74	10.90	10.43	1	False	-0.13
59293	0.76	0.80	10.17	9.63	1	False	-0.02
59294	0.77	0.78	10.04	9.91	1	False	-0.71
60278	0.79	0.81	9.77	9.49	1	False	-1.18
60313	0.81	0.85	9.49	8.85	1	False	0.22
60918	0.87	0.90	8.49	7.87	1	False	-0.05
60919	0.90	0.94	7.87	6.78	1	False	1.80
60974	0.76	0.80	10.17	9.63	1	False	-0.02
61046	0.62	0.66	11.78	11.35	1	False	-0.14
61065	0.54	0.64	12.60	11.57	1	False	2.59
76750	0.31	0.31	14.98	14.98	1	False	-1.24
76752	0.85	0.86	8.85	8.68	1	False	-0.62
76759	0.83	0.86	9.18	8.68	1	False	-0.39
76765	0.89	0.91	8.09	7.64	1	False	-0.83
76769	0.53	0.56	12.70	12.40	1	False	-0.46
76772	0.85	0.87	8.85	8.49	1	False	-1.08
76774	0.38	0.37	14.22	14.33	1	False	-0.62
76781	0.69	0.67	11.02	11.24	1	False	0.59
76782	0.58	0.61	12.19	11.88	1	False	-0.55
76784	0.45	0.45	13.50	13.50	1	False	-0.91
76795	0.84	0.85	9.02	8.85	1	False	-0.63
76836	0.72	0.67	10.67	11.24	1	False	2.13
76840	0.51	0.57	12.90	12.29	1	False	0.86
76841	0.36	0.37	14.43	14.33	1	False	-0.91
76843	0.49	0.47	13.10	13.30	1	False	0.03
76859	0.77	0.80	10.04	9.63	1	False	-0.60
76864	0.54	0.56	12.60	12.40	1	False	-0.91
76879	0.47	0.51	13.30	12.90	1	False	0.09
76895	0.65	0.65	11.46	11.46	1	False	-0.46
76899	0.36	0.39	14.43	14.17	4	False	-0.24
76899	0.39	0.39	14.17	14.17	4	False	-1.06
76904	0.81	0.84	9.49	9.02	1	False	-0.49
76906	0.67	0.67	11.24	11.24	1	False	-0.41
76911	0.55	0.55	12.50	12.50	1	False	-0.69
76913	0.46	0.38	13.40	14.22	1	False	2.58
76915	0.65	0.61	11.46	11.88	1	False	1.33
76917	0.89	0.88	8.09	8.30	1	False	1.16
76971	0.85	0.83	8.85	9.18	1	False	1.51
76979	0.61	0.66	11.88	11.35	1	False	0.33
77003	0.92	0.91	7.38	7.64	1	False	1.53
77019	0.77	0.77	10.04	10.04	1	False	-0.14

Table I-2. 2009–10 MontCAS: Delta Analysis Results – Mathematics Grade 4

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
242880	0.70	0.68	10.90	11.13	1	False	0.52
242889	0.49	0.64	13.10	11.57	1	True	4.25
243037	0.72	0.79	10.67	9.77	1	False	1.44
243174	0.74	0.70	10.43	10.90	1	False	1.51
243174	0.70	0.70	10.90	10.90	1	False	-0.29
244321	0.88	0.90	8.30	7.87	1	False	-0.74
244388	0.69	0.70	11.02	10.90	1	False	-0.72
248102	0.48	0.48	13.20	13.20	1	False	-0.79
43167	0.67	0.68	11.24	11.13	1	False	-0.76
43173	0.75	0.74	10.30	10.43	1	False	0.29
43276	0.78	0.80	9.91	9.63	1	False	-0.91
43298	0.53	0.54	12.70	12.60	1	False	-0.93
43312	0.83	0.85	9.18	8.85	1	False	-0.89
43314	0.77	0.76	10.04	10.17	1	False	0.36
43330	0.90	0.94	7.87	6.78	1	False	1.53
43332	0.61	0.52	11.88	12.80	1	False	2.75
43369	0.60	0.66	11.99	11.35	1	False	0.82
61780	0.55	0.56	12.50	12.40	1	False	-0.97
61804	0.42	0.45	13.81	13.50	1	False	0.04
62214	0.84	0.84	9.02	9.02	1	False	0.13
62320	0.86	0.88	8.68	8.30	1	False	-0.83
62381	0.81	0.80	9.49	9.63	1	False	0.54
62389	0.71	0.70	10.79	10.90	1	False	0.15
62401	0.76	0.79	10.17	9.77	1	False	-0.42
62486	0.50	0.52	13.00	12.82	4	False	-0.59
76763	0.49	0.54	13.10	12.60	1	False	0.59
76788	0.42	0.42	13.81	13.81	1	False	-0.93
76794	0.63	0.65	11.67	11.46	1	False	-0.75
76812	0.59	0.61	12.09	11.88	1	False	-0.68
76819	0.71	0.69	10.79	11.02	1	False	0.56
76823	0.62	0.62	11.78	11.78	1	False	-0.48
76827	0.58	0.56	12.19	12.40	1	False	0.15
76834	0.36	0.37	14.43	14.33	1	False	-0.52
76844	0.47	0.44	13.30	13.60	1	False	0.26
76856	0.70	0.67	10.90	11.24	1	False	0.91
76888	0.64	0.64	11.57	11.57	1	False	-0.43
76892	0.55	0.57	12.50	12.29	1	False	-0.61
76924	0.53	0.53	12.70	12.70	1	False	-0.68
76933	0.56	0.56	12.40	12.40	1	False	-0.62
76935	0.77	0.79	10.04	9.77	1	False	-0.91
76939	0.51	0.56	12.90	12.40	1	False	0.55
76941	0.69	0.75	11.02	10.30	1	False	0.88
76959	0.79	0.82	9.77	9.34	1	False	-0.38
76961	0.69	0.71	11.02	10.79	1	False	-0.84
76963	0.45	0.42	13.50	13.81	1	False	0.22
76965	0.55	0.58	12.50	12.19	1	False	-0.25
76995	0.77	0.81	10.04	9.49	1	False	0.11
77022	0.68	0.69	11.13	11.02	1	False	-0.74
77050	0.64	0.64	11.57	11.57	1	False	-0.43
77063	0.50	0.53	12.97	12.75	4	False	-0.42
77063	0.49	0.53	13.08	12.75	4	False	-0.04

Table I-3. 2009–10 MontCAS: Delta Analysis Results – Mathematics Grade 5

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
242955	0.43	0.47	13.71	13.30	1	False	0.07
242957	0.63	0.64	11.67	11.62	4	False	-1.01
43409	0.35	0.32	14.54	14.87	1	False	0.42
43413	0.56	0.59	12.40	12.09	1	False	-0.42
43421	0.75	0.74	10.30	10.43	1	False	-0.19
43471	0.68	0.70	11.13	10.90	1	False	-0.82
43477	0.65	0.67	11.46	11.24	1	False	-0.84
43514	0.43	0.43	13.71	13.71	1	False	-0.91
59800	0.64	0.65	11.57	11.46	1	False	-1.23
59856	0.64	0.58	11.57	12.19	1	False	1.83
60370	0.64	0.66	11.57	11.35	1	False	-0.84
60371	0.83	0.81	9.18	9.49	1	False	0.63
60417	0.68	0.68	11.13	11.13	1	False	-0.76
60504	0.74	0.79	10.43	9.77	1	False	0.91
60971	0.59	0.60	12.09	11.99	1	False	-1.25
61052	0.37	0.39	14.35	14.09	4	False	-0.48
61052	0.36	0.39	14.49	14.09	4	False	0.08
62025	0.66	0.62	11.35	11.78	1	False	1.02
77163	0.66	0.72	11.35	10.67	1	False	1.09
77179	0.66	0.68	11.35	11.13	1	False	-0.83
77181	0.66	0.62	11.35	11.78	1	False	1.02
77186	0.65	0.66	11.46	11.35	1	False	-1.23
77191	0.74	0.69	10.43	11.02	1	False	1.75
77193	0.79	0.81	9.77	9.49	1	False	-0.66
77198	0.42	0.46	13.81	13.40	1	False	0.08
77204	0.61	0.64	11.88	11.57	1	False	-0.40
77205	0.78	0.78	9.91	9.91	1	False	-0.68
77207	0.56	0.59	12.40	12.09	1	False	-0.42
77208	0.40	0.40	14.01	14.01	1	False	-0.93
77210	0.63	0.69	11.67	11.02	1	False	1.00
77217	0.74	0.71	10.43	10.79	1	False	0.79
77219	0.37	0.35	14.33	14.54	1	False	-0.05
77220	0.64	0.63	11.57	11.67	1	False	-0.34
77243	0.57	0.59	12.29	12.09	1	False	-0.85
77247	0.69	0.71	11.02	10.79	1	False	-0.82
77257	0.67	0.68	11.24	11.13	1	False	-1.23
77259	0.83	0.88	9.18	8.30	1	False	1.80
77265	0.38	0.47	14.22	13.30	1	False	2.26
77270	0.69	0.69	11.02	11.02	1	False	-0.75
77274	0.80	0.79	9.63	9.77	1	False	-0.08
77295	0.83	0.87	9.18	8.49	1	False	0.99
77298	0.34	0.41	14.65	13.91	1	False	1.53
77310	0.54	0.53	12.60	12.70	1	False	-0.42
77314	0.51	0.54	12.90	12.60	1	False	-0.41
77318	0.44	0.46	13.60	13.40	1	False	-0.78
77321	0.34	0.33	14.65	14.76	1	False	-0.51
77325	0.82	0.80	9.34	9.63	1	False	0.58
77330	0.37	0.30	14.33	15.10	1	False	2.27

Table I-4. 2009–10 MontCAS: Delta Analysis Results – Mathematics Grade 6

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
43921	0.49	0.56	13.10	12.40	1	False	0.49
43944	0.60	0.67	11.99	11.24	1	False	0.65
43949	0.44	0.48	13.60	13.20	1	False	-0.42
43956	0.58	0.63	12.19	11.67	1	False	-0.03
43963	0.69	0.70	11.02	10.90	1	False	-1.12
43992	0.53	0.52	12.70	12.80	1	False	-0.44
43995	0.78	0.80	9.91	9.63	1	False	-0.69
44004	0.74	0.77	10.43	10.04	1	False	-0.39
44088	0.80	0.85	9.63	8.85	1	False	0.81
44088	0.86	0.85	8.68	8.85	1	False	-0.32
60885	0.60	0.59	11.99	12.09	1	False	-0.45
60901	0.51	0.52	12.90	12.80	1	False	-1.03
61156	0.53	0.47	12.70	13.30	1	False	1.06
62029	0.60	0.64	11.99	11.57	1	False	-0.32
62050	0.43	0.38	13.71	14.22	1	False	0.83
62060	0.81	0.85	9.49	8.85	1	False	0.38
62994	0.41	0.45	13.91	13.50	1	False	-0.41
63005	0.43	0.41	13.71	13.91	1	False	-0.10
77317	0.48	0.47	13.20	13.30	1	False	-0.42
77320	0.82	0.81	9.34	9.49	1	False	-0.38
77323	0.69	0.77	11.02	10.04	1	False	1.35
77340	0.77	0.62	10.04	11.78	1	True	4.36
77376	0.62	0.63	11.78	11.67	1	False	-1.08
77377	0.95	0.93	6.42	7.10	1	False	1.11
77378	0.58	0.51	12.19	12.90	1	False	1.36
77380	0.56	0.54	12.40	12.60	1	False	-0.14
77398	0.34	0.40	14.65	14.01	1	False	0.25
77449	0.49	0.55	13.10	12.50	1	False	0.19
77451	0.48	0.49	13.20	13.10	1	False	-1.02
77476	0.75	0.81	10.30	9.49	1	False	0.90
77502	0.62	0.63	11.78	11.67	1	False	-1.08
77509	0.62	0.69	11.78	11.02	1	False	0.70
77515	0.48	0.50	13.20	13.00	1	False	-1.01
77522	0.60	0.64	11.99	11.57	1	False	-0.32
77538	0.72	0.69	10.67	11.02	1	False	0.25
77553	0.56	0.53	12.40	12.70	1	False	0.16
77582	0.65	0.66	11.46	11.35	1	False	-1.09
77608	0.40	0.46	14.01	13.40	1	False	0.19
77614	0.76	0.75	10.17	10.30	1	False	-0.43
77621	0.71	0.76	10.79	10.17	1	False	0.28
77625	0.87	0.87	8.49	8.49	1	False	-0.85
77630	0.56	0.57	12.40	12.29	1	False	-1.05
77633	0.62	0.60	11.78	11.99	1	False	-0.14
77642	0.50	0.49	13.00	13.10	1	False	-0.43
77649	0.28	0.25	15.30	15.70	4	False	0.52
242549	0.35	0.47	14.54	13.30	1	False	2.05
243201	0.46	0.47	13.40	13.30	1	False	-1.02
43863	0.44	0.46	13.60	13.40	1	False	-1.02
43912	0.48	0.51	13.20	12.90	1	False	-0.71

Table I-5. 2009–10 MontCAS: Delta Analysis Results – Mathematics Grade 7

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
43700	0.52	0.51	12.80	12.90	1	False	-0.35
43721	0.54	0.56	12.60	12.40	1	False	-0.49
43782	0.50	0.49	13.00	13.10	1	False	-0.35
43787	0.65	0.62	11.46	11.78	1	False	0.85
43799	0.73	0.73	10.55	10.55	1	False	-0.99
43836	0.62	0.60	11.78	11.99	1	False	0.23
43922	0.42	0.43	13.81	13.73	4	False	-1.24
43922	0.43	0.43	13.73	13.73	4	False	-0.90
61195	0.37	0.36	14.33	14.43	1	False	-0.28
61204	0.49	0.49	13.10	13.10	1	False	-0.92
61205	0.36	0.38	14.43	14.22	1	False	-0.49
61206	0.79	0.78	9.77	9.91	1	False	-0.23
61228	0.43	0.42	13.71	13.81	1	False	-0.32
61250	0.71	0.72	10.79	10.67	1	False	-0.92
61279	0.88	0.87	8.30	8.49	1	False	0.05
61742	0.52	0.54	12.80	12.60	1	False	-0.51
61745	0.64	0.66	11.57	11.35	1	False	-0.39
61766	0.88	0.88	8.30	8.30	1	False	-1.05
62948	0.38	0.37	14.22	14.33	1	False	-0.29
86280	0.78	0.78	9.91	9.91	1	False	-1.01
86295	0.69	0.67	11.02	11.24	1	False	0.30
86296	0.57	0.56	12.29	12.40	1	False	-0.36
86297	0.68	0.74	11.13	10.43	1	False	2.39
86300	0.33	0.33	14.76	14.76	1	False	-0.87
86302	0.53	0.51	12.70	12.90	1	False	0.21
86305	0.40	0.40	14.01	14.01	1	False	-0.89
86311	0.51	0.46	12.90	13.40	1	False	1.93
86313	0.60	0.59	11.99	12.09	1	False	-0.36
86348	0.62	0.63	11.78	11.67	1	False	-1.02
86349	0.50	0.55	13.00	12.50	1	False	1.20
86366	0.55	0.56	12.50	12.40	1	False	-1.07
86369	0.65	0.68	11.46	11.13	1	False	0.26
86374	0.62	0.60	11.78	11.99	1	False	0.23
86381	0.39	0.45	14.12	13.50	1	False	1.81
86382	0.77	0.79	10.04	9.77	1	False	-0.03
86431	0.42	0.42	13.81	13.81	1	False	-0.90
86438	0.94	0.95	6.78	6.42	1	False	0.57
86448	0.51	0.53	12.90	12.70	1	False	-0.51
86453	0.37	0.37	14.33	14.33	1	False	-0.88
86455	0.53	0.60	12.70	11.99	1	False	2.41
86458	0.49	0.55	13.10	12.50	1	False	1.77
86473	0.63	0.66	11.67	11.35	1	False	0.22
86482	0.76	0.72	10.17	10.67	1	False	1.81
86535	0.80	0.79	9.63	9.77	1	False	-0.21
86549	0.53	0.52	12.70	12.80	1	False	-0.36
86555	0.73	0.75	10.55	10.30	1	False	-0.18
86568	0.50	0.49	13.00	13.10	1	False	-0.35
86591	0.86	0.84	8.68	9.02	1	False	0.91
86615	0.40	0.42	14.01	13.81	1	False	-0.51
86622	0.57	0.57	12.29	12.29	1	False	-0.94
86631	0.45	0.50	13.50	13.00	1	False	1.19
86650	0.71	0.77	10.79	10.04	1	False	2.63
86675	0.44	0.44	13.60	13.60	1	False	-0.90
86683	0.76	0.74	10.17	10.43	1	False	0.43
86689	0.64	0.63	11.57	11.67	1	False	-0.35

Table I-6. 2009–10 MontCAS: Delta Analysis Results – Mathematics Grade 8

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
244538	0.88	0.88	8.30	8.30	1	False	0.79
44123	0.47	0.50	13.30	13.00	1	False	-0.23
44137	0.58	0.60	12.19	11.99	1	False	-0.99
44141	0.77	0.80	10.04	9.63	1	False	-1.02
44160	0.49	0.46	13.10	13.40	1	False	0.14
44176	0.33	0.34	14.76	14.65	1	False	-0.40
44199	0.52	0.52	12.80	12.80	1	False	-0.86
44205	0.77	0.75	10.04	10.30	1	False	1.10
44220	0.57	0.55	12.29	12.50	1	False	0.07
62943	0.51	0.57	12.90	12.29	1	False	0.74
63025	0.43	0.45	13.71	13.50	1	False	-0.44
63148	0.31	0.27	14.98	15.45	1	False	0.06
63203	0.41	0.43	13.91	13.71	1	False	-0.36
63219	0.55	0.56	12.50	12.40	1	False	-1.12
63252	0.51	0.46	12.90	13.40	1	False	0.95
63287	0.62	0.67	11.78	11.24	1	False	0.08
87527	0.64	0.65	11.57	11.46	1	False	-0.80
87580	0.68	0.66	11.13	11.35	1	False	0.57
87583	0.65	0.68	11.46	11.13	1	False	-0.80
87593	0.74	0.77	10.43	10.04	1	False	-0.99
87598	0.93	0.97	7.10	5.48	1	False	2.35
87606	0.53	0.55	12.70	12.50	1	False	-0.82
87623	0.36	0.32	14.43	14.87	1	False	0.15
87658	0.50	0.53	13.00	12.70	1	False	-0.34
87661	0.27	0.31	15.45	14.98	1	False	1.17
87799	0.53	0.54	12.70	12.60	1	False	-1.19
87802	0.62	0.64	11.78	11.57	1	False	-1.12
87808	0.36	0.38	14.43	14.22	1	False	-0.14
87834	0.50	0.57	13.00	12.29	1	False	1.15
87841	0.47	0.50	13.30	13.00	1	False	-0.23
88019	0.60	0.63	11.99	11.67	1	False	-0.66
88086	0.52	0.52	12.80	12.80	1	False	-0.86
88174	0.58	0.55	12.19	12.50	1	False	0.49
88189	0.28	0.31	15.33	14.98	1	False	0.68
88363	0.54	0.43	12.60	13.71	1	True	3.29
88838	0.64	0.64	11.57	11.57	1	False	-0.41

Table I-7. 2009–10 MontCAS: Delta Analysis Results – Mathematics Grade 10

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
242987	0.71	0.74	10.79	10.43	1	False	-0.39
243118	0.40	0.38	14.01	14.22	1	False	-0.14
243153	0.43	0.46	13.71	13.40	1	False	0.12
248846	0.26	0.25	15.57	15.70	1	False	-1.08
43743	0.34	0.27	14.65	15.45	1	True	3.19
43822	0.61	0.66	11.88	11.35	1	False	0.95
59377	0.53	0.53	12.70	12.70	1	False	-1.00
59397	0.70	0.72	10.90	10.67	1	False	-1.10
61265	0.37	0.37	14.33	14.33	1	False	-1.46
61281	0.71	0.69	10.79	11.02	1	False	0.92
61298	0.31	0.31	14.98	14.98	1	False	-1.31
61319	0.43	0.44	13.71	13.60	1	False	-1.08

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
61324	0.89	0.88	8.09	8.30	1	False	1.55
62177	0.78	0.75	9.91	10.30	1	False	2.12
62286	0.34	0.31	14.65	14.98	1	False	0.42
62292	0.38	0.39	14.22	14.12	1	False	-0.91
62333	0.59	0.63	12.09	11.67	1	False	0.33
62368	0.55	0.60	12.50	11.99	1	False	1.00
77352	0.20	0.22	16.37	16.09	1	False	0.73
77354	0.60	0.60	11.99	11.99	1	False	-0.79
77368	0.57	0.59	12.29	12.09	1	False	-0.87
77370	0.54	0.58	12.60	12.19	1	False	0.41
77371	0.77	0.79	10.04	9.77	1	False	-1.13
77382	0.37	0.35	14.33	14.54	1	False	-0.20
77384	0.31	0.31	14.98	14.98	1	False	-1.31
77392	0.52	0.56	12.80	12.40	1	False	0.45
77394	0.39	0.36	14.12	14.43	1	False	0.47
77415	0.51	0.53	12.90	12.70	1	False	-0.72
77428	0.61	0.65	11.88	11.46	1	False	0.31
77432	0.63	0.68	11.67	11.13	1	False	0.95
77480	0.48	0.51	13.20	12.90	1	False	-0.04
77484	0.69	0.69	11.02	11.02	1	False	-0.51
77485	0.56	0.55	12.40	12.50	1	False	-0.31
77507	0.43	0.44	13.71	13.60	1	False	-1.08
77551	0.50	0.53	13.00	12.70	1	False	-0.10
77561	0.54	0.54	12.60	12.60	1	False	-0.97
77562	0.35	0.31	14.54	14.98	1	False	1.09
77570	0.47	0.45	13.30	13.50	1	False	0.03
77571	0.37	0.40	14.33	14.01	1	False	0.36
77612	0.57	0.60	12.29	11.99	1	False	-0.26
77618	0.76	0.76	10.17	10.17	1	False	-0.27
77619	0.48	0.45	13.20	13.50	1	False	0.65
77623	0.54	0.59	12.60	12.09	1	False	1.01

Table I-8. 2009–10 MontCAS: Delta Analysis Results – Reading Grade 3

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
68808	0.80	0.85	9.63	8.85	1	False	0.37
68809	0.67	0.77	11.24	10.04	1	False	2.80
68810	0.50	0.54	13.00	12.60	1	False	0.37
68811	0.91	0.95	7.64	6.42	1	False	1.24
68812	0.54	0.53	12.60	12.70	1	False	-0.41
68814	0.53	0.48	12.70	13.20	1	False	1.17
68818	0.63	0.66	11.67	11.35	1	False	-0.55
92670	0.91	0.91	7.64	7.64	1	False	1.44
92670	0.89	0.91	8.09	7.64	1	False	-0.62
92673	0.61	0.59	11.88	12.09	1	False	0.35
92673	0.58	0.59	12.19	12.09	1	False	-1.05
92674	0.51	0.57	12.90	12.29	1	False	1.15
92674	0.52	0.57	12.80	12.29	1	False	0.70
92675	0.83	0.85	9.18	8.85	1	False	-0.60
92675	0.85	0.85	8.85	8.85	1	False	0.89
92677	0.61	0.63	11.88	11.67	1	False	-0.91
92695	0.79	0.80	9.77	9.63	1	False	-0.10
92695	0.80	0.80	9.63	9.63	1	False	0.53
92696	0.60	0.56	11.99	12.40	1	False	1.12

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
92696	0.61	0.56	11.88	12.40	1	False	1.59
92739	0.78	0.81	9.91	9.49	1	False	-0.95
92739	0.80	0.81	9.63	9.49	1	False	-0.06
92742	0.68	0.69	11.13	11.02	1	False	-0.61
92742	0.67	0.69	11.24	11.02	1	False	-1.11
92742	0.67	0.69	11.24	11.02	1	False	-1.11
92742	0.68	0.69	11.13	11.02	1	False	-0.61
92743	0.67	0.69	11.24	11.02	1	False	-1.11
92743	0.63	0.69	11.67	11.02	1	False	0.80
92745	0.51	0.54	12.90	12.60	1	False	-0.08
92745	0.54	0.54	12.60	12.60	1	False	-0.82
92746	0.88	0.89	8.30	8.09	1	False	0.30
92746	0.85	0.89	8.85	8.09	1	False	-0.06
92748	0.85	0.87	8.85	8.49	1	False	-0.57
92748	0.83	0.87	9.18	8.49	1	False	-0.20
92748	0.83	0.87	9.18	8.49	1	False	-0.20
92748	0.85	0.87	8.85	8.49	1	False	-0.57
92749	0.80	0.79	9.63	9.77	1	False	1.10
92749	0.78	0.79	9.91	9.77	1	False	-0.15
92750	0.73	0.75	10.55	10.30	1	False	-0.89
92750	0.78	0.75	9.91	10.30	1	False	1.99
92751	0.74	0.75	10.43	10.30	1	False	-0.33
92751	0.72	0.75	10.67	10.30	1	False	-0.83
92752	0.62	0.61	11.78	11.88	1	False	-0.02
92752	0.62	0.61	11.78	11.88	1	False	-0.02
92755	0.81	0.82	9.49	9.34	1	False	-0.01
92755	0.81	0.82	9.49	9.34	1	False	-0.01
92755	0.81	0.82	9.49	9.34	1	False	-0.01
92755	0.77	0.82	10.04	9.34	1	False	0.26
92758	0.79	0.80	9.77	9.63	1	False	-0.10
92758	0.78	0.80	9.91	9.63	1	False	-0.72
92758	0.79	0.80	9.77	9.63	1	False	-0.10
92758	0.78	0.80	9.91	9.63	1	False	-0.72
92761	0.38	0.33	14.22	14.76	4	False	0.63
92761	0.37	0.33	14.33	14.76	4	False	0.15
92763	0.67	0.73	11.24	10.55	1	False	0.75
92763	0.72	0.73	10.67	10.55	1	False	-0.43
92763	0.71	0.73	10.79	10.55	1	False	-0.96
92763	0.72	0.73	10.67	10.55	1	False	-0.43
92765	0.57	0.55	12.29	12.50	1	False	0.14
92765	0.56	0.55	12.40	12.50	1	False	-0.32
92766	0.58	0.69	12.19	11.02	1	True	3.15
92766	0.58	0.69	12.19	11.02	1	True	3.15
92767	0.80	0.82	9.63	9.34	1	False	-0.67
92768	0.80	0.84	9.63	9.02	1	False	-0.31
92771	0.55	0.61	12.50	11.88	1	False	1.01
92771	0.58	0.61	12.19	11.88	1	False	-0.37
92773	0.71	0.72	10.79	10.67	1	False	-0.47
92773	0.71	0.72	10.79	10.67	1	False	-0.47
92775	0.85	0.88	8.85	8.30	1	False	-0.90
92775	0.87	0.88	8.49	8.30	1	False	0.26
92775	0.84	0.88	9.02	8.30	1	False	-0.14
92775	0.87	0.88	8.49	8.30	1	False	0.26
92777	0.57	0.57	12.29	12.29	1	False	-0.68
92778	0.64	0.68	11.57	11.13	1	False	-0.14
92778	0.63	0.68	11.67	11.13	1	False	0.34

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
92779	0.55	0.57	12.50	12.29	1	False	-0.66
92781	0.58	0.57	12.19	12.29	1	False	-0.22
92781	0.57	0.57	12.29	12.29	1	False	-0.68
92783	0.31	0.30	14.96	15.16	4	False	-1.08
92783	0.30	0.30	15.07	15.16	4	False	-0.67
92786	0.59	0.61	12.09	11.88	1	False	-0.83
92786	0.55	0.61	12.50	11.88	1	False	1.01
92789	0.89	0.92	8.09	7.38	1	False	-0.59
92789	0.89	0.92	8.09	7.38	1	False	-0.59
92791	0.72	0.84	10.67	9.02	1	True	4.37
92791	0.76	0.84	10.17	9.02	1	True	2.14
92792	0.78	0.81	9.91	9.49	1	False	-0.95
92792	0.74	0.81	10.43	9.49	1	False	1.38
92794	0.71	0.77	10.79	10.04	1	False	0.75
92794	0.74	0.77	10.43	10.04	1	False	-0.88
92795	0.60	0.60	11.99	11.99	1	False	-0.54
92795	0.62	0.60	11.78	11.99	1	False	0.40
92795	0.59	0.60	12.09	11.99	1	False	-1.01
92795	0.58	0.60	12.19	11.99	1	False	-0.79
92797	0.73	0.76	10.55	10.17	1	False	-0.86
92797	0.74	0.76	10.43	10.17	1	False	-0.85
92798	0.82	0.85	9.34	8.85	1	False	-0.96
92798	0.82	0.85	9.34	8.85	1	False	-0.96
92800	0.75	0.75	10.30	10.30	1	False	0.23
92800	0.75	0.75	10.30	10.30	1	False	0.23
92802	0.53	0.51	12.70	12.90	1	False	-0.05
92802	0.52	0.51	12.80	12.90	1	False	-0.50
92803	0.74	0.77	10.43	10.04	1	False	-0.88
92803	0.73	0.77	10.55	10.04	1	False	-0.33
92808	0.67	0.69	11.24	11.02	1	False	-1.11
92808	0.69	0.69	11.02	11.02	1	False	-0.10

Table I-9. 2009–10 MontCAS: Delta Analysis Results – Reading Grade 4

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
67330	0.70	0.77	10.90	10.04	1	False	0.72
67330	0.72	0.77	10.67	10.04	1	False	-0.17
67333	0.62	0.63	11.78	11.67	1	False	-1.02
67334	0.56	0.56	12.40	12.40	1	False	-0.75
67346	0.73	0.75	10.55	10.30	1	False	-1.31
67350	0.75	0.76	10.30	10.17	1	False	-0.83
67354	0.56	0.52	12.40	12.80	1	False	0.71
67359	0.58	0.59	12.19	12.09	1	False	-1.09
67365	0.67	0.73	11.24	10.55	1	False	0.18
67367	0.57	0.67	12.29	11.24	1	False	1.69
67368	0.65	0.68	11.46	11.13	1	False	-1.09
67371	0.76	0.74	10.17	10.43	1	False	0.57
67374	0.70	0.68	10.90	11.13	1	False	0.34
67382	0.35	0.35	14.51	14.51	4	False	-1.14
93837	0.76	0.82	10.17	9.34	1	False	0.51
93837	0.79	0.82	9.77	9.34	1	False	-1.01
93837	0.79	0.82	9.77	9.34	1	False	-1.01
93837	0.74	0.82	10.43	9.34	1	False	1.47
93838	0.66	0.71	11.35	10.79	1	False	-0.26

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
93838	0.67	0.71	11.24	10.79	1	False	-0.68
93838	0.62	0.71	11.78	10.79	1	False	1.37
93838	0.62	0.71	11.78	10.79	1	False	1.37
93842	0.57	0.55	12.29	12.50	1	False	0.00
93842	0.55	0.55	12.50	12.50	1	False	-0.77
93846	0.64	0.67	11.57	11.24	1	False	-1.09
93846	0.58	0.67	12.19	11.24	1	False	1.30
93849	0.70	0.69	10.90	11.02	1	False	-0.06
93849	0.69	0.69	11.02	11.02	1	False	-0.50
93855	0.89	0.87	8.09	8.49	1	False	1.49
93855	0.90	0.87	7.87	8.49	1	False	2.33
93857	0.71	0.67	10.79	11.24	1	False	1.19
93857	0.70	0.67	10.90	11.24	1	False	0.75
93862	0.75	0.75	10.30	10.30	1	False	-0.37
93862	0.75	0.75	10.30	10.30	1	False	-0.37
93863	0.76	0.76	10.17	10.17	1	False	-0.35
93863	0.78	0.76	9.91	10.17	1	False	0.66
93874	0.69	0.68	11.02	11.13	1	False	-0.09
93874	0.70	0.68	10.90	11.13	1	False	0.34
93874	0.65	0.68	11.46	11.13	1	False	-1.09
93874	0.73	0.68	10.55	11.13	1	False	1.69
93876	0.78	0.83	9.91	9.18	1	False	0.07
93881	0.66	0.67	11.35	11.24	1	False	-0.96
93881	0.73	0.67	10.55	11.24	1	False	2.10
93888	0.53	0.55	12.70	12.50	1	False	-1.33
93888	0.46	0.55	13.40	12.50	1	False	1.35
93895	0.81	0.87	9.49	8.49	1	False	0.96
93939	0.61	0.69	11.88	11.02	1	False	0.93
93939	0.63	0.69	11.67	11.02	1	False	0.13
93944	0.74	0.80	10.43	9.63	1	False	0.40
93944	0.72	0.80	10.67	9.63	1	False	1.32
93946	0.81	0.80	9.49	9.63	1	False	0.31
93948	0.70	0.69	10.90	11.02	1	False	-0.06
93951	0.76	0.83	10.17	9.18	1	False	1.08
94002	0.76	0.81	10.17	9.49	1	False	-0.03
94002	0.71	0.81	10.79	9.49	1	False	2.30
94004	0.70	0.77	10.90	10.04	1	False	0.72
94004	0.73	0.77	10.55	10.04	1	False	-0.62
94015	0.81	0.81	9.49	9.49	1	False	-0.22
94015	0.79	0.81	9.77	9.49	1	False	-1.31
94016	0.61	0.64	11.88	11.57	1	False	-1.06
94016	0.64	0.64	11.57	11.57	1	False	-0.60
94027	0.65	0.69	11.46	11.02	1	False	-0.68
94027	0.64	0.69	11.57	11.02	1	False	-0.27
94048	0.72	0.74	10.67	10.43	1	False	-1.31
94048	0.72	0.74	10.67	10.43	1	False	-1.31
94050	0.72	0.72	10.67	10.67	1	False	-0.44
94050	0.71	0.72	10.79	10.67	1	False	-0.89
94072	0.67	0.75	11.24	10.30	1	False	1.08
94072	0.67	0.75	11.24	10.30	1	False	1.08
94077	0.79	0.77	9.77	10.04	1	False	0.71
94077	0.78	0.77	9.91	10.04	1	False	0.19
94079	0.86	0.88	8.68	8.30	1	False	-1.42
94079	0.86	0.88	8.68	8.30	1	False	-1.42
94083	0.52	0.58	12.80	12.19	1	False	0.16
94083	0.51	0.58	12.90	12.19	1	False	0.54

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
94092	0.83	0.86	9.18	8.68	1	False	-0.87
94092	0.83	0.86	9.18	8.68	1	False	-0.87
94095	0.72	0.69	10.67	11.02	1	False	0.83
94095	0.73	0.69	10.55	11.02	1	False	1.28
94108	0.75	0.78	10.30	9.91	1	False	-1.08
94108	0.77	0.78	10.04	9.91	1	False	-0.81
94111	0.76	0.76	10.17	10.17	1	False	-0.35
94111	0.76	0.76	10.17	10.17	1	False	-0.35
94116	0.45	0.46	13.50	13.40	1	False	-1.32
94116	0.43	0.46	13.71	13.40	1	False	-0.77
94120	0.72	0.75	10.67	10.30	1	False	-1.10
94120	0.73	0.75	10.55	10.30	1	False	-1.31
94120	0.71	0.75	10.79	10.30	1	False	-0.65
94120	0.75	0.75	10.30	10.30	1	False	-0.37
94139	0.42	0.37	13.78	14.38	4	False	1.17
94139	0.43	0.37	13.71	14.38	4	False	1.46

Table I-10. 2009–10 MontCAS: Delta Analysis Results – Reading Grade 5

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
65312	0.78	0.77	9.91	10.04	1	False	0.40
65317	0.71	0.73	10.79	10.55	1	False	-0.92
65368	0.68	0.70	11.13	10.90	1	False	-0.81
65379	0.83	0.82	9.18	9.34	1	False	0.80
65383	0.77	0.82	10.04	9.34	1	False	0.57
65387	0.51	0.41	12.90	13.91	1	False	2.52
69235	0.79	0.80	9.77	9.63	1	False	-0.59
93353	0.57	0.59	12.29	12.09	1	False	-0.41
93353	0.58	0.59	12.19	12.09	1	False	-0.84
93366	0.72	0.73	10.67	10.55	1	False	-0.89
93366	0.73	0.73	10.55	10.55	1	False	-0.38
93375	0.87	0.90	8.49	7.87	1	False	-0.41
93375	0.87	0.90	8.49	7.87	1	False	-0.41
93378	0.68	0.76	11.13	10.17	1	False	1.99
93378	0.70	0.76	10.90	10.17	1	False	1.02
93381	0.80	0.86	9.63	8.68	1	False	1.36
93381	0.83	0.86	9.18	8.68	1	False	-0.57
93385	0.55	0.59	12.50	12.09	1	False	0.46
93385	0.55	0.59	12.50	12.09	1	False	0.46
93389	0.61	0.64	11.88	11.57	1	False	-0.15
93389	0.62	0.64	11.78	11.57	1	False	-0.60
93414	0.80	0.82	9.63	9.34	1	False	-1.13
93414	0.78	0.82	9.91	9.34	1	False	0.00
93415	0.76	0.79	10.17	9.77	1	False	-0.55
93415	0.77	0.79	10.04	9.77	1	False	-1.10
93416	0.74	0.76	10.43	10.17	1	False	-1.01
93416	0.78	0.76	9.91	10.17	1	False	0.90
93419	0.88	0.90	8.30	7.87	1	False	-1.07
93419	0.88	0.90	8.30	7.87	1	False	-1.07
93419	0.89	0.90	8.09	7.87	1	False	-0.19
93419	0.88	0.90	8.30	7.87	1	False	-1.07
93420	0.77	0.81	10.04	9.49	1	False	0.00
93420	0.75	0.81	10.30	9.49	1	False	1.10
93421	0.79	0.81	9.77	9.49	1	False	-1.15

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
93421	0.79	0.81	9.77	9.49	1	False	-1.15
93428	0.83	0.84	9.18	9.02	1	False	-0.42
93428	0.80	0.84	9.63	9.02	1	False	0.04
93431	0.64	0.64	11.57	11.57	1	False	-0.81
93431	0.68	0.64	11.13	11.57	1	False	1.06
93444	0.69	0.75	11.02	10.30	1	False	1.02
93444	0.68	0.75	11.13	10.30	1	False	1.50
93444	0.63	0.75	11.67	10.30	1	True	3.82
93444	0.63	0.75	11.67	10.30	1	True	3.82
93446	0.52	0.49	12.80	13.10	1	False	-0.17
93446	0.55	0.49	12.50	13.10	1	False	1.12
93448	0.74	0.78	10.43	9.91	1	False	0.00
93448	0.75	0.78	10.30	9.91	1	False	-0.53
93451	0.54	0.49	12.60	13.10	1	False	0.69
93451	0.56	0.49	12.40	13.10	1	False	1.55
93462	0.49	0.49	13.10	13.10	4	False	-0.85
93462	0.48	0.49	13.18	13.10	4	False	-0.53
93510	0.67	0.69	11.24	11.02	1	False	-0.78
93510	0.67	0.69	11.24	11.02	1	False	-0.78
93520	0.67	0.68	11.24	11.13	1	False	-1.10
93520	0.69	0.68	11.02	11.13	1	False	-0.14
93524	0.72	0.71	10.67	10.79	1	False	0.02
93524	0.75	0.71	10.30	10.79	1	False	1.59
93529	0.80	0.79	9.63	9.77	1	False	0.55
93529	0.78	0.79	9.91	9.77	1	False	-0.64
93533	0.64	0.58	11.57	12.19	1	False	1.60
93533	0.63	0.58	11.67	12.19	1	False	1.15
93536	0.74	0.76	10.43	10.17	1	False	-1.01
93536	0.76	0.76	10.17	10.17	1	False	-0.22
93537	0.59	0.63	12.09	11.67	1	False	0.33
93537	0.62	0.63	11.78	11.67	1	False	-1.01
93601	0.77	0.79	10.04	9.77	1	False	-1.10
93601	0.78	0.79	9.91	9.77	1	False	-0.64
93601	0.79	0.79	9.77	9.77	1	False	-0.05
93601	0.80	0.79	9.63	9.77	1	False	0.55
93611	0.66	0.67	11.35	11.24	1	False	-1.14
93611	0.67	0.67	11.24	11.24	1	False	-0.67
93612	0.64	0.63	11.57	11.67	1	False	-0.40
93612	0.62	0.63	11.78	11.67	1	False	-1.01
93615	0.81	0.85	9.49	8.85	1	False	0.06
93615	0.85	0.85	8.85	8.85	1	False	0.34
93616	0.69	0.72	11.02	10.67	1	False	-0.39
93616	0.69	0.72	11.02	10.67	1	False	-0.39
93623	0.73	0.75	10.55	10.30	1	False	-0.98
93623	0.74	0.75	10.43	10.30	1	False	-0.81
93623	0.74	0.75	10.43	10.30	1	False	-0.81
93623	0.76	0.75	10.17	10.30	1	False	0.27
93626	0.64	0.61	11.57	11.88	1	False	0.41
93626	0.63	0.61	11.67	11.88	1	False	-0.04
93628	0.70	0.68	10.90	11.13	1	False	0.34
93628	0.72	0.68	10.67	11.13	1	False	1.34
93631	0.82	0.82	9.34	9.34	1	False	0.13
93631	0.83	0.82	9.18	9.34	1	False	0.80
93635	0.71	0.69	10.79	11.02	1	False	0.41
93635	0.70	0.69	10.90	11.02	1	False	-0.09
93638	0.67	0.69	11.24	11.02	1	False	-0.78

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
93638	0.66	0.69	11.35	11.02	1	False	-0.31
93639	0.88	0.90	8.30	7.87	1	False	-1.07
93639	0.90	0.90	7.87	7.87	1	False	0.75
93668	0.41	0.48	13.91	13.25	4	False	2.03
93668	0.43	0.48	13.68	13.25	4	False	1.04
93695	0.64	0.69	11.57	11.02	1	False	0.61
93695	0.66	0.69	11.35	11.02	1	False	-0.31
93698	0.79	0.82	9.77	9.34	1	False	-0.58
93698	0.81	0.82	9.49	9.34	1	False	-0.51
93700	0.87	0.90	8.49	7.87	1	False	-0.41
93700	0.88	0.90	8.30	7.87	1	False	-1.07
93700	0.87	0.90	8.49	7.87	1	False	-0.41
93700	0.86	0.90	8.68	7.87	1	False	0.38
93704	0.77	0.79	10.04	9.77	1	False	-1.10
93704	0.77	0.79	10.04	9.77	1	False	-1.10
93705	0.55	0.54	12.50	12.60	1	False	-0.81
93705	0.60	0.54	11.99	12.60	1	False	1.37
93709	0.57	0.62	12.29	11.78	1	False	0.79
93709	0.58	0.62	12.19	11.78	1	False	0.36
93711	0.82	0.83	9.34	9.18	1	False	-0.47
93711	0.82	0.83	9.34	9.18	1	False	-0.47

Table I-11. 2009–10 MontCAS: Delta Analysis Results – Reading Grade 6

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
95033	0.81	0.81	9.49	9.49	1	False	-0.20
95033	0.84	0.81	9.02	9.49	1	False	1.46
95036	0.81	0.85	9.49	8.85	1	False	-0.33
95036	0.80	0.85	9.63	8.85	1	False	0.19
95041	0.52	0.52	12.80	12.80	1	False	-1.01
95041	0.53	0.52	12.70	12.80	1	False	-1.17
95045	0.81	0.79	9.49	9.77	1	False	0.70
95045	0.82	0.79	9.34	9.77	1	False	1.24
95077	0.49	0.55	13.10	12.50	1	False	1.02
95077	0.46	0.55	13.40	12.50	1	False	2.09
95085	0.89	0.91	8.09	7.64	1	False	-1.09
95085	0.89	0.91	8.09	7.64	1	False	-1.09
95088	0.86	0.89	8.68	8.09	1	False	-0.81
95088	0.81	0.89	9.49	8.09	1	False	2.08
95089	0.66	0.62	11.35	11.78	1	False	0.41
95089	0.68	0.62	11.13	11.78	1	False	1.20
95092	0.83	0.84	9.18	9.02	1	False	-0.59
95092	0.82	0.84	9.34	9.02	1	False	-1.14
95093	0.80	0.80	9.63	9.63	1	False	-0.26
95093	0.79	0.80	9.77	9.63	1	False	-0.76
95145	0.68	0.66	11.13	11.35	1	False	-0.16
95157	0.81	0.83	9.49	9.18	1	False	-1.17
95157	0.79	0.83	9.77	9.18	1	False	-0.35
95171	0.49	0.53	13.13	12.67	4	False	0.55
95171	0.48	0.53	13.20	12.67	4	False	0.82
95183	0.45	0.40	13.50	14.01	1	False	-0.19
95183	0.47	0.40	13.30	14.01	1	False	0.53
95202	0.65	0.63	11.46	11.67	1	False	-0.31
95202	0.66	0.63	11.35	11.67	1	False	0.08

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
95218	0.70	0.70	10.90	10.90	1	False	-0.77
95218	0.71	0.70	10.79	10.90	1	False	-0.35
95228	0.78	0.80	9.91	9.63	1	False	-1.25
95228	0.78	0.80	9.91	9.63	1	False	-1.25
95101	0.76	0.86	10.17	8.68	1	False	2.68
95101	0.80	0.86	9.63	8.68	1	False	0.75
95105	0.51	0.61	12.90	11.88	1	False	2.25
95105	0.50	0.61	13.00	11.88	1	False	2.60
95110	0.77	0.77	10.04	10.04	1	False	-0.42
95110	0.75	0.77	10.30	10.04	1	False	-1.20
95114	0.77	0.76	10.04	10.17	1	False	-0.01
95114	0.78	0.76	9.91	10.17	1	False	0.46
95115	0.65	0.72	11.46	10.67	1	False	0.95
95115	0.68	0.72	11.13	10.67	1	False	-0.22
95115	0.70	0.72	10.90	10.67	1	False	-1.03
95115	0.71	0.72	10.79	10.67	1	False	-1.09
95121	0.50	0.60	13.00	11.99	1	False	2.28
95121	0.55	0.60	12.50	11.99	1	False	0.48
95132	0.65	0.64	11.46	11.57	1	False	-0.65
95132	0.65	0.64	11.46	11.57	1	False	-0.65
95145	0.67	0.66	11.24	11.35	1	False	-0.55
95231	0.74	0.75	10.43	10.30	1	False	-0.97
95231	0.75	0.75	10.30	10.30	1	False	-0.53
95289	0.58	0.53	12.19	12.70	1	False	0.32
95289	0.57	0.53	12.29	12.70	1	False	-0.04
95299	0.54	0.49	12.60	13.10	1	False	0.15
95299	0.54	0.49	12.60	13.10	1	False	0.15
95305	0.65	0.63	11.46	11.67	1	False	-0.31
95305	0.67	0.63	11.24	11.67	1	False	0.47
95330	0.67	0.72	11.24	10.67	1	False	0.17
95330	0.72	0.72	10.67	10.67	1	False	-0.67
95335	0.68	0.68	11.13	11.13	1	False	-0.86
95335	0.70	0.68	10.90	11.13	1	False	-0.05
95342	0.81	0.77	9.49	10.04	1	False	1.56
95342	0.77	0.77	10.04	10.04	1	False	-0.42
95348	0.71	0.68	10.79	11.13	1	False	0.37
95348	0.71	0.68	10.79	11.13	1	False	0.37
95351	0.70	0.76	10.90	10.17	1	False	0.53
95351	0.72	0.76	10.67	10.17	1	False	-0.30
95353	0.69	0.62	11.02	11.78	1	False	1.60
95353	0.66	0.62	11.35	11.78	1	False	0.41
95358	0.77	0.84	10.04	9.02	1	False	1.12
95358	0.82	0.84	9.34	9.02	1	False	-1.14
95363	0.79	0.88	9.77	8.30	1	False	2.45
95363	0.87	0.88	8.49	8.30	1	False	-0.42
95368	0.88	0.90	8.30	7.87	1	False	-1.08
95368	0.83	0.90	9.18	7.87	1	False	1.69
95369	0.79	0.83	9.77	9.18	1	False	-0.35
95369	0.72	0.83	10.67	9.18	1	False	2.84
95371	0.82	0.84	9.34	9.02	1	False	-1.14
95371	0.85	0.84	8.85	9.02	1	False	0.58
95375	0.72	0.76	10.67	10.17	1	False	-0.30
95375	0.70	0.76	10.90	10.17	1	False	0.53
95381	0.62	0.61	11.78	11.88	1	False	-0.78
95381	0.60	0.61	11.99	11.88	1	False	-1.01
95386	0.56	0.53	12.40	12.70	1	False	-0.40

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
95386	0.54	0.53	12.60	12.70	1	False	-1.12
95387	0.69	0.66	11.02	11.35	1	False	0.24
95387	0.67	0.66	11.24	11.35	1	False	-0.55
95393	0.64	0.64	11.57	11.57	1	False	-1.03
95393	0.60	0.64	11.99	11.57	1	False	-0.01
95397	0.51	0.50	12.95	13.05	4	False	-1.27
95397	0.50	0.50	12.97	13.05	4	False	-1.18
95398	0.75	0.82	10.30	9.34	1	False	1.04
95398	0.76	0.82	10.17	9.34	1	False	0.59
95409	0.79	0.80	9.77	9.63	1	False	-0.76
95409	0.79	0.80	9.77	9.63	1	False	-0.76
95410	0.69	0.74	11.02	10.43	1	False	0.14
95410	0.69	0.74	11.02	10.43	1	False	0.14
95421	0.85	0.85	8.85	8.85	1	False	0.05
95421	0.85	0.85	8.85	8.85	1	False	0.05
95431	0.66	0.65	11.35	11.46	1	False	-0.60
95431	0.65	0.65	11.46	11.46	1	False	-0.99
95435	0.63	0.65	11.67	11.46	1	False	-0.79
95435	0.64	0.65	11.57	11.46	1	False	-1.17
95445	0.57	0.58	12.29	12.19	1	False	-0.89
95445	0.60	0.58	11.99	12.19	1	False	-0.55
95450	0.78	0.77	9.91	10.04	1	False	0.05
95450	0.76	0.77	10.17	10.04	1	False	-0.89
97773	0.75	0.73	10.30	10.55	1	False	0.26
97773	0.76	0.73	10.17	10.55	1	False	0.71

Table I-12. 2009–10 MontCAS: Delta Analysis Results – Reading Grade 7

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
41859	0.54	0.60	12.60	11.99	1	False	2.09
41860	0.72	0.74	10.67	10.43	1	False	-0.99
41864	0.81	0.85	9.49	8.85	1	False	-0.21
41866	0.72	0.76	10.67	10.17	1	False	0.09
41867	0.78	0.80	9.91	9.63	1	False	-1.42
41868	0.82	0.85	9.34	8.85	1	False	-0.97
68493	0.85	0.83	8.85	9.18	1	False	1.91
68495	0.60	0.60	11.99	11.99	1	False	-1.02
68497	0.75	0.75	10.30	10.30	1	False	-0.62
68498	0.79	0.81	9.77	9.49	1	False	-1.45
68507	0.84	0.88	9.02	8.30	1	False	-0.19
68510	0.81	0.83	9.49	9.18	1	False	-1.31
68510	0.79	0.83	9.77	9.18	1	False	-0.18
68514	0.70	0.73	10.90	10.55	1	False	-0.33
68610	0.84	0.90	9.02	7.87	1	False	1.65
68611	0.60	0.66	11.99	11.35	1	False	1.72
68612	0.72	0.79	10.67	9.77	1	False	1.82
68613	0.68	0.63	11.13	11.67	1	False	1.09
68614	0.53	0.56	12.70	12.40	1	False	0.83
68616	0.75	0.79	10.30	9.77	1	False	-0.04
68620	0.84	0.88	9.02	8.30	1	False	-0.19
92341	0.65	0.63	11.46	11.67	1	False	-0.59
92341	0.65	0.63	11.46	11.67	1	False	-0.59
92342	0.67	0.64	11.24	11.57	1	False	0.06
92342	0.70	0.64	10.90	11.57	1	False	1.78

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
92342	0.70	0.64	10.90	11.57	1	False	1.78
92342	0.70	0.64	10.90	11.57	1	False	1.78
92343	0.76	0.75	10.17	10.30	1	False	0.03
92343	0.74	0.75	10.43	10.30	1	False	-1.25
92345	0.74	0.76	10.43	10.17	1	False	-1.14
92345	0.72	0.76	10.67	10.17	1	False	0.09
92347	0.69	0.65	11.02	11.46	1	False	0.74
92347	0.68	0.65	11.13	11.46	1	False	0.16
92348	0.65	0.69	11.46	11.02	1	False	0.48
92348	0.66	0.69	11.35	11.02	1	False	-0.07
92350	0.76	0.75	10.17	10.30	1	False	0.03
92350	0.72	0.75	10.67	10.30	1	False	-0.46
92531	0.62	0.61	11.78	11.88	1	False	-1.31
92531	0.62	0.61	11.78	11.88	1	False	-1.31
92535	0.74	0.76	10.43	10.17	1	False	-1.14
92536	0.66	0.64	11.35	11.57	1	False	-0.50
92540	0.58	0.56	12.19	12.40	1	False	-1.20
92540	0.60	0.56	11.99	12.40	1	False	-0.15
92541	0.65	0.63	11.46	11.67	1	False	-0.59
92541	0.63	0.63	11.67	11.67	1	False	-1.26
92543	0.73	0.70	10.55	10.90	1	False	0.71
92543	0.75	0.70	10.30	10.90	1	False	1.97
92545	0.63	0.62	11.67	11.78	1	False	-1.22
92545	0.59	0.62	12.09	11.78	1	False	0.40
92549	0.84	0.84	9.02	9.02	1	False	0.37
92549	0.84	0.84	9.02	9.02	1	False	0.37
92554	0.82	0.81	9.34	9.49	1	False	0.77
92554	0.84	0.81	9.02	9.49	1	False	2.38
92555	0.69	0.67	11.02	11.24	1	False	-0.21
92555	0.70	0.67	10.90	11.24	1	False	0.37
92558	0.58	0.53	12.19	12.70	1	False	0.11
92558	0.57	0.53	12.29	12.70	1	False	-0.41
92559	0.56	0.55	12.40	12.50	1	False	-1.14
92562	0.50	0.47	12.97	13.33	4	False	-1.16
92562	0.51	0.47	12.87	13.33	4	False	-0.65
92567	0.75	0.81	10.30	9.49	1	False	1.19
92567	0.79	0.81	9.77	9.49	1	False	-1.45
92583	0.65	0.67	11.46	11.24	1	False	-0.49
92583	0.62	0.67	11.78	11.24	1	False	1.14
92588	0.69	0.73	11.02	10.55	1	False	0.25
92588	0.68	0.73	11.13	10.55	1	False	0.82
92589	0.71	0.71	10.79	10.79	1	False	-0.99
92589	0.66	0.71	11.35	10.79	1	False	0.92
92591	0.84	0.88	9.02	8.30	1	False	-0.19
92591	0.86	0.88	8.68	8.30	1	False	-1.00
92593	0.75	0.76	10.30	10.17	1	False	-1.17
92593	0.75	0.76	10.30	10.17	1	False	-1.17
92596	0.63	0.58	11.67	12.19	1	False	0.57
92596	0.62	0.58	11.78	12.19	1	False	0.03
92600	0.82	0.82	9.34	9.34	1	False	0.12
92600	0.82	0.82	9.34	9.34	1	False	0.12
92605	0.74	0.79	10.43	9.77	1	False	0.59
92605	0.73	0.79	10.55	9.77	1	False	1.21
92606	0.60	0.59	11.99	12.09	1	False	-1.47
92606	0.56	0.59	12.40	12.09	1	False	0.61
92608	0.76	0.82	10.17	9.34	1	False	1.19

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
92608	0.78	0.82	9.91	9.34	1	False	-0.15
92608	0.78	0.82	9.91	9.34	1	False	-0.15
92608	0.82	0.82	9.34	9.34	1	False	0.12
92611	0.53	0.52	12.70	12.77	4	False	-0.80
92611	0.52	0.52	12.82	12.77	4	False	-0.16

Table I-13. 2009–10 MontCAS: Delta Analysis Results – Reading Grade 8

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
67937	0.68	0.70	11.13	10.90	1	False	-1.17
67938	0.78	0.78	9.91	9.91	1	False	-0.20
67944	0.78	0.81	9.91	9.49	1	False	-0.61
67948	0.66	0.63	11.35	11.67	1	False	0.95
67952	0.89	0.88	8.09	8.30	1	False	0.73
67953	0.54	0.54	12.60	12.60	1	False	-0.40
67966	0.82	0.83	9.34	9.18	1	False	-0.76
68470	0.68	0.69	11.13	11.02	1	False	-0.73
68470	0.68	0.69	11.13	11.02	1	False	-0.73
68473	0.77	0.74	10.04	10.43	1	False	1.27
68473	0.74	0.74	10.43	10.43	1	False	-0.24
68475	0.75	0.78	10.30	9.91	1	False	-0.71
68475	0.76	0.78	10.17	9.91	1	False	-1.21
68475	0.78	0.78	9.91	9.91	1	False	-0.20
68475	0.77	0.78	10.04	9.91	1	False	-0.73
68476	0.79	0.77	9.77	10.04	1	False	0.86
68476	0.78	0.77	9.91	10.04	1	False	0.32
68478	0.78	0.82	9.91	9.34	1	False	-0.03
68478	0.79	0.82	9.77	9.34	1	False	-0.57
68480	0.69	0.68	11.02	11.13	1	False	0.16
68480	0.67	0.68	11.24	11.13	1	False	-0.73
68480	0.67	0.68	11.24	11.13	1	False	-0.73
68480	0.65	0.68	11.46	11.13	1	False	-0.86
68487	0.67	0.66	11.24	11.35	1	False	0.13
68487	0.68	0.66	11.13	11.35	1	False	0.57
68499	0.78	0.77	9.91	10.04	1	False	0.32
68499	0.78	0.77	9.91	10.04	1	False	0.32
68500	0.81	0.80	9.49	9.63	1	False	0.39
68500	0.78	0.80	9.91	9.63	1	False	-1.18
68501	0.73	0.73	10.55	10.55	1	False	-0.25
68501	0.72	0.73	10.67	10.55	1	False	-0.72
68504	0.73	0.75	10.55	10.30	1	False	-1.20
68504	0.73	0.75	10.55	10.30	1	False	-1.20
68508	0.63	0.63	11.67	11.67	1	False	-0.33
68508	0.63	0.63	11.67	11.67	1	False	-0.33
68511	0.58	0.52	12.19	12.85	4	False	2.19
68511	0.56	0.52	12.45	12.85	4	False	1.18
68698	0.65	0.69	11.46	11.02	1	False	-0.43
68699	0.71	0.77	10.79	10.04	1	False	0.69
68702	0.66	0.67	11.35	11.24	1	False	-0.73
68714	0.77	0.72	10.04	10.67	1	False	2.21
68724	0.80	0.79	9.63	9.77	1	False	0.37
68725	0.81	0.80	9.49	9.63	1	False	0.39
68726	0.83	0.85	9.18	8.85	1	False	-1.03
95604	0.80	0.79	9.63	9.77	1	False	0.37

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
95604	0.78	0.79	9.91	9.77	1	False	-0.73
95627	0.78	0.83	9.91	9.18	1	False	0.57
95627	0.80	0.83	9.63	9.18	1	False	-0.53
95637	0.70	0.80	10.90	9.63	1	False	2.74
95637	0.72	0.80	10.67	9.63	1	False	1.82
95644	0.69	0.75	11.02	10.30	1	False	0.60
95644	0.67	0.75	11.24	10.30	1	False	1.49
95647	0.71	0.69	10.79	11.02	1	False	0.63
95647	0.71	0.69	10.79	11.02	1	False	0.63
95649	0.58	0.66	12.19	11.35	1	False	1.18
95649	0.54	0.66	12.60	11.35	1	False	2.79
95656	0.75	0.78	10.30	9.91	1	False	-0.71
95656	0.75	0.78	10.30	9.91	1	False	-0.71
95656	0.77	0.78	10.04	9.91	1	False	-0.73
95656	0.76	0.78	10.17	9.91	1	False	-1.21
95688	0.72	0.73	10.67	10.55	1	False	-0.72
95688	0.70	0.73	10.90	10.55	1	False	-0.81
95691	0.73	0.77	10.55	10.04	1	False	-0.25
95691	0.76	0.77	10.17	10.04	1	False	-0.72
95700	0.66	0.67	11.35	11.24	1	False	-0.73
95700	0.65	0.67	11.46	11.24	1	False	-1.16
95703	0.69	0.71	11.02	10.79	1	False	-1.17
95703	0.70	0.71	10.90	10.79	1	False	-0.72
95706	0.83	0.83	9.18	9.18	1	False	-0.15
95706	0.84	0.83	9.02	9.18	1	False	0.49
95708	0.67	0.70	11.24	10.90	1	False	-0.85
95708	0.69	0.70	11.02	10.90	1	False	-0.72
95838	0.70	0.74	10.90	10.43	1	False	-0.34
95838	0.70	0.74	10.90	10.43	1	False	-0.34
95843	0.74	0.75	10.43	10.30	1	False	-0.72
95843	0.73	0.75	10.55	10.30	1	False	-1.20
95844	0.55	0.62	12.50	11.78	1	False	0.73
95844	0.52	0.62	12.80	11.78	1	False	1.92
95845	0.74	0.78	10.43	9.91	1	False	-0.22
95845	0.73	0.78	10.55	9.91	1	False	0.27
95847	0.70	0.73	10.90	10.55	1	False	-0.81
95847	0.70	0.73	10.90	10.55	1	False	-0.81
95851	0.80	0.84	9.63	9.02	1	False	0.10
95851	0.84	0.84	9.02	9.02	1	False	-0.13
95853	0.76	0.79	10.17	9.77	1	False	-0.68
95853	0.78	0.79	9.91	9.77	1	False	-0.73
95855	0.65	0.70	11.46	10.90	1	False	0.02
95855	0.67	0.70	11.24	10.90	1	False	-0.85
95856	0.80	0.84	9.63	9.02	1	False	0.10
95856	0.81	0.84	9.49	9.02	1	False	-0.48
95863	0.78	0.87	9.91	8.49	1	True	3.25
95863	0.80	0.87	9.63	8.49	1	True	2.15
95866	0.66	0.65	11.35	11.46	1	False	0.12
95866	0.68	0.65	11.13	11.46	1	False	0.99
95867	0.78	0.84	9.91	9.02	1	False	1.20
95867	0.82	0.84	9.34	9.02	1	False	-1.07
95869	0.48	0.57	13.18	12.29	4	False	1.40
95869	0.51	0.57	12.92	12.29	4	False	0.41

Table I-14. 2009–10 MontCAS: Delta Analysis Results – Reading Grade 10

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
67599	0.86	0.87	8.68	8.49	1	False	-0.50
67687	0.77	0.83	10.04	9.18	1	False	1.16
67720	0.54	0.54	12.60	12.60	1	False	-0.85
67741	0.68	0.62	11.13	11.78	1	False	1.46
67741	0.61	0.62	11.88	11.78	1	False	-0.78
67750	0.76	0.79	10.17	9.77	1	False	-0.45
67753	0.80	0.78	9.63	9.91	1	False	0.76
67753	0.81	0.78	9.49	9.91	1	False	1.35
67757	0.68	0.74	11.13	10.43	1	False	1.06
94892	0.67	0.71	11.24	10.79	1	False	0.20
94892	0.69	0.71	11.02	10.79	1	False	-0.70
94894	0.67	0.62	11.24	11.78	1	False	1.01
94894	0.62	0.62	11.78	11.78	1	False	-1.16
94898	0.63	0.65	11.67	11.46	1	False	-0.48
94898	0.64	0.65	11.57	11.46	1	False	-0.91
94903	0.74	0.72	10.43	10.67	1	False	0.29
94903	0.74	0.72	10.43	10.67	1	False	0.29
94912	0.69	0.71	11.02	10.79	1	False	-0.70
94912	0.71	0.71	10.79	10.79	1	False	-0.74
94924	0.60	0.64	11.99	11.57	1	False	0.40
94924	0.62	0.64	11.78	11.57	1	False	-0.44
94929	0.77	0.79	10.04	9.77	1	False	-0.98
94929	0.76	0.79	10.17	9.77	1	False	-0.45
94931	0.80	0.81	9.63	9.49	1	False	-0.77
94931	0.83	0.81	9.18	9.49	1	False	1.06
94941	0.82	0.83	9.34	9.18	1	False	-0.68
94941	0.81	0.83	9.49	9.18	1	False	-1.09
94941	0.80	0.83	9.63	9.18	1	False	-0.50
94941	0.82	0.83	9.34	9.18	1	False	-0.68
94943	0.80	0.81	9.63	9.49	1	False	-0.77
94943	0.81	0.81	9.49	9.49	1	False	-0.18
94943	0.81	0.81	9.49	9.49	1	False	-0.18
94943	0.79	0.81	9.77	9.49	1	False	-1.04
94955	0.47	0.46	13.33	13.38	4	False	-0.72
94955	0.46	0.46	13.40	13.38	4	False	-0.42
95026	0.77	0.81	10.04	9.49	1	False	0.06
95026	0.77	0.81	10.04	9.49	1	False	0.06
95026	0.77	0.81	10.04	9.49	1	False	0.06
95026	0.79	0.81	9.77	9.49	1	False	-1.04
95030	0.82	0.86	9.34	8.68	1	False	0.13
95030	0.82	0.86	9.34	8.68	1	False	0.13
95138	0.67	0.72	11.24	10.67	1	False	0.63
95138	0.68	0.72	11.13	10.67	1	False	0.18
95154	0.80	0.86	9.63	8.68	1	False	1.32
95154	0.80	0.86	9.63	8.68	1	False	1.32
95164	0.76	0.78	10.17	9.91	1	False	-0.94
95164	0.76	0.78	10.17	9.91	1	False	-0.94
95187	0.73	0.61	10.55	11.88	1	True	4.19
95187	0.73	0.61	10.55	11.88	1	True	4.19
95207	0.89	0.93	8.09	7.10	1	False	0.82
95207	0.87	0.93	8.49	7.10	1	False	2.44
95216	0.77	0.75	10.04	10.30	1	False	0.51
95216	0.79	0.75	9.77	10.30	1	False	1.61
95234	0.75	0.74	10.30	10.43	1	False	-0.08
95234	0.72	0.74	10.67	10.43	1	False	-0.81

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
95273	0.67	0.63	11.24	11.67	1	False	0.63
95273	0.67	0.63	11.24	11.67	1	False	0.63
95279	0.64	0.64	11.57	11.57	1	False	-1.07
95279	0.63	0.64	11.67	11.57	1	False	-0.87
95285	0.64	0.67	11.57	11.24	1	False	-0.12
95285	0.67	0.67	11.24	11.24	1	False	-0.93
95290	0.56	0.59	12.40	12.09	1	False	0.17
95290	0.60	0.59	11.99	12.09	1	False	-0.88
95293	0.81	0.75	9.49	10.30	1	False	2.77
95293	0.79	0.75	9.77	10.30	1	False	1.61
95338	0.80	0.77	9.63	10.04	1	False	1.25
95338	0.78	0.77	9.91	10.04	1	False	0.12
95340	0.76	0.76	10.17	10.17	1	False	-0.47
95340	0.75	0.76	10.30	10.17	1	False	-0.99
95361	0.73	0.76	10.55	10.17	1	False	-0.38
95361	0.76	0.76	10.17	10.17	1	False	-0.47
95367	0.72	0.71	10.67	10.79	1	False	-0.26
95367	0.72	0.71	10.67	10.79	1	False	-0.26
95374	0.63	0.65	11.67	11.46	1	False	-0.48
95374	0.62	0.65	11.78	11.46	1	False	-0.05
95377	0.87	0.88	8.49	8.30	1	False	-0.45
95377	0.88	0.88	8.30	8.30	1	False	0.33
95391	0.65	0.67	11.46	11.24	1	False	-0.55
95391	0.65	0.67	11.46	11.24	1	False	-0.55
94842	0.64	0.65	11.57	11.46	1	False	-0.91
94842	0.65	0.65	11.46	11.46	1	False	-1.03
94843	0.83	0.87	9.18	8.49	1	False	0.17
94843	0.82	0.87	9.34	8.49	1	False	0.80
94848	0.75	0.78	10.30	9.91	1	False	-0.43
94848	0.73	0.78	10.55	9.91	1	False	0.57
94852	0.87	0.87	8.49	8.49	1	False	0.25
94852	0.86	0.87	8.68	8.49	1	False	-0.50
94854	0.86	0.87	8.68	8.49	1	False	-0.50
94854	0.86	0.87	8.68	8.49	1	False	-0.50
94857	0.88	0.91	8.30	7.64	1	False	-0.31
94857	0.88	0.91	8.30	7.64	1	False	-0.31
94861	0.72	0.76	10.67	10.17	1	False	0.10
94861	0.73	0.76	10.55	10.17	1	False	-0.38
94863	0.68	0.70	11.13	10.90	1	False	-0.67
94863	0.70	0.70	10.90	10.90	1	False	-0.79
94874	0.72	0.73	10.67	10.55	1	False	-1.12
94874	0.69	0.73	11.02	10.55	1	False	0.16
94877	0.76	0.78	10.17	9.91	1	False	-0.94
94877	0.77	0.78	10.04	9.91	1	False	-0.90
94879	0.61	0.69	11.88	11.02	1	False	1.97
94879	0.63	0.69	11.67	11.02	1	False	1.12
94882	0.60	0.55	11.99	12.50	1	False	0.60
94882	0.58	0.55	12.19	12.50	1	False	-0.24
94887	0.52	0.54	12.77	12.60	4	False	-0.14
94887	0.53	0.54	12.75	12.60	4	False	-0.24
94889	0.82	0.83	9.34	9.18	1	False	-0.68
94889	0.82	0.83	9.34	9.18	1	False	-0.68
94890	0.60	0.63	11.99	11.67	1	False	0.02
94890	0.60	0.63	11.99	11.67	1	False	0.02

Table I-15. 2009–10 MontCAS: Delta Analysis Results – Science Grade 4

IREF	*Old mean*	*New mean*	*Old delta*	*New delta*	*Maximum*	*Discard*	*Standardized difference*
53393	0.63	0.61	11.67	11.88	1	False	0.57
53393	0.65	0.61	11.46	11.88	1	False	1.80
53659	0.88	0.88	8.30	8.30	1	False	0.52
53659	0.87	0.88	8.49	8.30	1	False	-0.60
55464	0.85	0.87	8.85	8.49	1	False	-1.13
55464	0.86	0.87	8.68	8.49	1	False	-0.60
55629	0.39	0.36	14.12	14.43	1	False	0.36
55629	0.39	0.36	14.12	14.43	1	False	0.36
56340	0.61	0.61	11.88	11.88	1	False	-0.64
56340	0.60	0.61	11.99	11.88	1	False	-1.24
57860	0.61	0.63	11.88	11.67	1	False	-0.96
57860	0.61	0.63	11.88	11.67	1	False	-0.96
57874	0.84	0.87	9.02	8.49	1	False	-0.16
57874	0.83	0.87	9.18	8.49	1	False	0.77
75408	0.49	0.54	13.10	12.60	1	False	1.02
75408	0.51	0.54	12.90	12.60	1	False	-0.13
75418	0.67	0.69	11.24	11.02	1	False	-1.09
75418	0.70	0.69	10.90	11.02	1	False	0.30
75421	0.55	0.51	12.50	12.90	1	False	1.35
75421	0.55	0.51	12.50	12.90	1	False	1.35
75423	0.85	0.84	8.85	9.02	1	False	1.25
75423	0.83	0.84	9.18	9.02	1	False	-0.64
75427	0.32	0.33	14.93	14.79	4	False	-0.35
75427	0.32	0.33	14.93	14.79	4	False	-0.35
75517	0.57	0.59	12.29	12.09	1	False	-0.85
75517	0.55	0.59	12.50	12.09	1	False	0.31
75690	0.50	0.57	13.00	12.29	1	False	2.09
75690	0.49	0.57	13.10	12.29	1	False	2.67
75694	0.68	0.64	11.13	11.57	1	False	1.98
75694	0.68	0.64	11.13	11.57	1	False	1.98
75702	0.87	0.89	8.49	8.09	1	False	-1.02
75702	0.86	0.89	8.68	8.09	1	False	0.04
75717	0.68	0.69	11.13	11.02	1	False	-1.01
75717	0.67	0.69	11.24	11.02	1	False	-1.09
75737	0.78	0.79	9.91	9.77	1	False	-0.74
75737	0.79	0.79	9.77	9.77	1	False	0.04
75741	0.65	0.66	11.46	11.35	1	False	-1.09
75741	0.65	0.66	11.46	11.35	1	False	-1.09
75782	0.43	0.46	13.71	13.40	1	False	0.14
75782	0.44	0.46	13.60	13.40	1	False	-0.44
75784	0.55	0.53	12.50	12.70	1	False	0.26
75784	0.52	0.53	12.80	12.70	1	False	-1.26
75788	0.85	0.89	8.85	8.09	1	False	1.05
75788	0.85	0.89	8.85	8.09	1	False	1.05
75801	0.55	0.57	12.50	12.29	1	False	-0.80
75801	0.53	0.57	12.70	12.29	1	False	0.36
75824	0.76	0.75	10.17	10.30	1	False	0.61
75824	0.77	0.75	10.04	10.30	1	False	1.35
75828	0.70	0.72	10.90	10.67	1	False	-1.15
75828	0.70	0.72	10.90	10.67	1	False	-1.15
75833	0.85	0.85	8.85	8.85	1	False	0.34
75833	0.85	0.85	8.85	8.85	1	False	0.34
75835	0.62	0.65	11.78	11.46	1	False	-0.40
75835	0.62	0.65	11.78	11.46	1	False	-0.40
75887	0.77	0.79	10.04	9.77	1	False	-1.23

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
75887	0.73	0.79	10.55	9.77	1	False	1.67
75889	0.47	0.51	13.30	12.90	1	False	0.54
75889	0.50	0.51	13.00	12.90	1	False	-1.19
75895	0.57	0.57	12.29	12.29	1	False	-0.77
75895	0.60	0.57	11.99	12.29	1	False	1.00
75899	0.70	0.72	10.90	10.67	1	False	-1.15
75899	0.71	0.72	10.79	10.67	1	False	-0.92
75901	0.64	0.67	11.57	11.24	1	False	-0.43
75901	0.65	0.67	11.46	11.24	1	False	-1.05
75902	0.73	0.72	10.55	10.67	1	False	0.44
75902	0.73	0.72	10.55	10.67	1	False	0.44
75908	0.73	0.73	10.55	10.55	1	False	-0.21
75908	0.72	0.73	10.67	10.55	1	False	-0.90
75910	0.82	0.87	9.34	8.49	1	False	1.66
75910	0.82	0.87	9.34	8.49	1	False	1.66
75912	0.71	0.71	10.79	10.79	1	False	-0.28
75912	0.71	0.71	10.79	10.79	1	False	-0.28
75923	0.51	0.51	12.90	12.90	1	False	-0.97
75923	0.50	0.51	13.00	12.90	1	False	-1.19
76285	0.80	0.82	9.63	9.34	1	False	-1.23
76285	0.77	0.82	10.04	9.34	1	False	1.14
76296	0.76	0.77	10.17	10.04	1	False	-0.79
76296	0.72	0.77	10.67	10.04	1	False	0.90
76394	0.73	0.76	10.55	10.17	1	False	-0.50
76394	0.71	0.76	10.79	10.17	1	False	0.87
76403	0.80	0.83	9.63	9.18	1	False	-0.38
76403	0.83	0.83	9.18	9.18	1	False	0.24

Table I-16. 2009–10 MontCAS: Delta Analysis Results – Science Grade 8

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
54454	0.59	0.63	12.09	11.67	1	False	-0.66
54454	0.60	0.63	11.99	11.67	1	False	-1.13
54543	0.60	0.60	11.99	11.99	1	False	-0.41
54543	0.58	0.60	12.19	11.99	1	False	-1.34
56805	0.83	0.84	9.18	9.02	1	False	0.30
56805	0.83	0.84	9.18	9.02	1	False	0.30
56828	0.43	0.44	13.71	13.60	1	False	-1.15
56828	0.42	0.44	13.81	13.60	1	False	-0.69
56833	0.45	0.46	13.50	13.40	1	False	-1.25
56833	0.44	0.46	13.60	13.40	1	False	-0.79
56851	0.38	0.43	14.22	13.71	1	False	0.77
56851	0.40	0.43	14.01	13.71	1	False	-0.17
56897	0.67	0.68	11.24	11.13	1	False	-0.49
56897	0.65	0.68	11.46	11.13	1	False	-1.32
56992	0.43	0.37	13.71	14.33	1	False	1.26
56992	0.43	0.37	13.71	14.33	1	False	1.26
89263	0.67	0.73	11.24	10.55	1	False	0.03
89263	0.65	0.73	11.46	10.55	1	False	1.01
89277	0.85	0.93	8.85	7.10	1	True	3.16
89277	0.84	0.93	9.02	7.10	1	True	3.91
89361	0.65	0.67	11.46	11.24	1	False	-1.03
89361	0.66	0.67	11.35	11.24	1	False	-0.54
89382	0.72	0.72	10.67	10.67	1	False	0.23

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
89382	0.73	0.72	10.55	10.67	1	False	0.77
89420	0.76	0.82	10.17	9.34	1	False	0.10
89420	0.78	0.82	9.91	9.34	1	False	-1.09
89444	0.39	0.38	14.12	14.22	1	False	-1.02
89444	0.37	0.38	14.33	14.22	1	False	-0.83
89452	0.66	0.70	11.35	10.90	1	False	-0.90
89452	0.68	0.70	11.13	10.90	1	False	-0.91
89457	0.44	0.45	13.60	13.50	1	False	-1.20
89457	0.45	0.45	13.50	13.50	1	False	-1.14
89468	0.53	0.48	12.70	13.20	1	False	1.26
89468	0.50	0.48	13.00	13.20	1	False	-0.09
89498	0.67	0.66	11.24	11.35	1	False	0.39
89498	0.68	0.66	11.13	11.35	1	False	0.89
89505	0.36	0.38	14.43	14.22	1	False	-0.35
89505	0.36	0.38	14.43	14.22	1	False	-0.35
89508	0.46	0.48	13.40	13.20	1	False	-0.89
89508	0.45	0.48	13.50	13.20	1	False	-0.44
89520	0.76	0.84	10.17	9.02	1	False	1.37
89520	0.81	0.84	9.49	9.02	1	False	-1.07
89582	0.31	0.35	14.98	14.54	1	False	0.84
89582	0.33	0.35	14.76	14.54	1	False	-0.17
89593	0.76	0.74	10.17	10.43	1	False	1.48
89593	0.73	0.74	10.55	10.43	1	False	-0.20
89634	0.48	0.47	13.20	13.30	1	False	-0.59
89634	0.46	0.47	13.40	13.30	1	False	-1.30
89639	0.51	0.47	12.90	13.30	1	False	0.76
89639	0.51	0.47	12.90	13.30	1	False	0.76
89647	0.69	0.75	11.02	10.30	1	False	0.01
89647	0.72	0.75	10.67	10.30	1	False	-1.24
89652	0.39	0.40	14.12	14.01	1	False	-0.94
89652	0.37	0.40	14.33	14.01	1	False	0.01
89691	0.52	0.57	12.80	12.29	1	False	0.03
89691	0.49	0.57	13.10	12.29	1	False	1.39
89693	0.63	0.61	11.67	11.88	1	False	0.59
89693	0.61	0.61	11.88	11.88	1	False	-0.36
89726	0.51	0.45	12.90	13.50	1	False	1.57
89726	0.48	0.45	13.20	13.50	1	False	0.22
89742	0.35	0.38	14.54	14.22	1	False	0.13
89742	0.32	0.38	14.87	14.22	1	False	1.62
89752	0.66	0.73	11.35	10.55	1	False	0.52
89752	0.67	0.73	11.24	10.55	1	False	0.03
89766	0.62	0.61	11.78	11.88	1	False	0.11
89766	0.66	0.61	11.35	11.88	1	False	2.04
89770	0.50	0.54	13.00	12.60	1	False	-0.28
89770	0.54	0.54	12.60	12.60	1	False	-0.70
89778	0.50	0.48	13.00	13.20	1	False	-0.09
89778	0.44	0.48	13.60	13.20	1	False	0.02
89795	0.68	0.71	11.13	10.79	1	False	-1.37
89795	0.70	0.71	10.90	10.79	1	False	-0.35
89817	0.70	0.69	10.90	11.02	1	False	0.57
89817	0.71	0.69	10.79	11.02	1	False	1.10
89849	0.60	0.65	11.99	11.46	1	False	-0.27
89849	0.58	0.65	12.19	11.46	1	False	0.66
89850	0.42	0.48	13.81	13.20	1	False	0.93
89850	0.44	0.48	13.60	13.20	1	False	0.02
89863	0.52	0.51	12.80	12.90	1	False	-0.40

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
89863	0.52	0.51	12.80	12.90	1	False	-0.40
89884	0.59	0.62	12.09	11.78	1	False	-1.08
89884	0.61	0.62	11.88	11.78	1	False	-0.78
89892	0.62	0.68	11.78	11.13	1	False	0.12
89892	0.64	0.68	11.57	11.13	1	False	-0.83
89911	0.64	0.64	11.57	11.57	1	False	-0.21
89911	0.65	0.64	11.46	11.57	1	False	0.28

Table I-17. 2009–10 MontCAS: Delta Analysis Results – Science Grade 10

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
52985	0.49	0.53	13.10	12.70	1	False	0.42
52985	0.50	0.53	13.00	12.70	1	False	0.03
53265	0.68	0.72	11.13	10.67	1	False	0.17
53265	0.71	0.72	10.79	10.67	1	False	-0.89
53812	0.49	0.50	13.10	13.00	1	False	-0.70
53812	0.49	0.50	13.10	13.00	1	False	-0.70
54221	0.62	0.67	11.78	11.24	1	False	0.62
54221	0.65	0.67	11.46	11.24	1	False	-0.65
55620	0.68	0.65	11.13	11.46	1	False	0.70
55620	0.63	0.65	11.67	11.46	1	False	-0.61
56234	0.47	0.48	13.30	13.20	1	False	-0.65
56234	0.47	0.48	13.30	13.20	1	False	-0.65
56702	0.58	0.56	12.19	12.40	1	False	-0.02
56702	0.56	0.56	12.40	12.40	1	False	-0.83
75433	0.53	0.51	12.70	12.90	1	False	-0.16
75433	0.52	0.51	12.80	12.90	1	False	-0.55
75440	0.48	0.48	13.20	13.20	1	False	-1.02
75440	0.46	0.48	13.40	13.20	1	False	-0.25
75445	0.35	0.38	14.54	14.22	1	False	0.46
75445	0.36	0.38	14.43	14.22	1	False	0.04
75456	0.55	0.55	12.50	12.50	1	False	-0.86
75456	0.54	0.55	12.60	12.50	1	False	-0.81
75611	0.56	0.62	12.40	11.78	1	False	1.06
75611	0.62	0.62	11.78	11.78	1	False	-0.68
75635	0.52	0.52	12.80	12.80	1	False	-0.93
75635	0.50	0.52	13.00	12.80	1	False	-0.35
75650	0.75	0.73	10.30	10.55	1	False	0.59
75650	0.75	0.73	10.30	10.55	1	False	0.59
75652	0.24	0.16	15.86	16.98	4	False	2.51
75652	0.24	0.16	15.86	16.98	4	False	2.51
75706	0.74	0.65	10.43	11.46	1	True	3.48
75706	0.75	0.65	10.30	11.46	1	True	3.97
75764	0.82	0.84	9.34	9.02	1	False	-0.79
75764	0.81	0.84	9.49	9.02	1	False	-0.20
75780	0.64	0.65	11.57	11.46	1	False	-1.03
75780	0.62	0.65	11.78	11.46	1	False	-0.20
75785	0.70	0.71	10.90	10.79	1	False	-0.91
75785	0.68	0.71	11.13	10.79	1	False	-0.26
75807	0.55	0.52	12.50	12.80	1	False	0.27
75807	0.56	0.52	12.40	12.80	1	False	0.67
75854	0.59	0.63	12.09	11.67	1	False	0.24
75854	0.58	0.63	12.19	11.67	1	False	0.65
75856	0.76	0.77	10.17	10.04	1	False	-0.79

continued

IREF	Old mean	New mean	Old delta	New delta	Maximum	Discard	Standardized difference
75856	0.76	0.77	10.17	10.04	1	False	-0.79
75859	0.60	0.58	11.99	12.19	1	False	0.03
75859	0.56	0.58	12.40	12.19	1	False	-0.48
75861	0.46	0.51	13.40	12.90	1	False	0.87
75861	0.46	0.51	13.40	12.90	1	False	0.87
75863	0.65	0.64	11.46	11.57	1	False	-0.21
75863	0.65	0.64	11.46	11.57	1	False	-0.21
75873	0.48	0.50	13.20	13.00	1	False	-0.30
75873	0.50	0.50	13.00	13.00	1	False	-0.97
75878	0.35	0.33	14.54	14.76	1	False	-0.53
75878	0.33	0.33	14.76	14.76	1	False	-0.68
75937	0.78	0.80	9.91	9.63	1	False	-0.80
75937	0.79	0.80	9.77	9.63	1	False	-0.73
75941	0.68	0.71	11.13	10.79	1	False	-0.26
75941	0.66	0.71	11.35	10.79	1	False	0.61
75948	0.49	0.49	13.10	13.10	1	False	-1.00
75948	0.48	0.49	13.20	13.10	1	False	-0.67
75958	0.39	0.42	14.12	13.81	1	False	0.32
75958	0.39	0.42	14.12	13.81	1	False	0.32
75970	0.57	0.54	12.29	12.60	1	False	0.32
75970	0.55	0.54	12.50	12.60	1	False	-0.48
75980	0.49	0.53	13.10	12.70	1	False	0.42
75980	0.48	0.53	13.20	12.70	1	False	0.82

Table I-18. 2009–10 MontCAS: Rescore Analysis Results by Subject and Grade

Subject	Grade	IREF	Maximum	Old mean	New mean	Old standard deviation	New standard deviation	Effect size	Discard
Mathematics	3	76899	4	1.62	1.60	1.06	1.02	-0.02	No
	4	77063	4	2.01	2.02	1.28	1.27	0.01	No
		62486	4	2.08	2.03	1.64	1.61	-0.03	No
	5	61052	4	1.48	1.52	1.35	1.32	0.03	No
		242957	4	2.58	2.50	1.33	1.27	-0.06	No
	6	77649	4	1.10	1.13	1.40	1.40	0.03	No
	7	43922	4	1.83	1.96	1.45	1.43	0.09	No
Reading	3	92761	4	1.46	1.18	0.81	0.83	-0.34	No
		92783	4	1.32	1.32	0.99	0.93	0.00	No
	4	67382	4	1.34	1.24	0.90	0.84	-0.11	No
		94139	4	1.81	1.56	0.73	0.84	-0.35	No
	5	93668	4	1.46	1.66	1.02	1.02	0.19	No
		93462	4	1.98	1.92	0.78	0.72	-0.08	No
	6	95397	4	1.99	1.83	1.06	1.07	-0.15	No
		95171	4	1.93	2.12	0.90	0.91	0.21	No
	7	92611	4	2.06	2.05	0.93	0.90	-0.01	No
		92562	4	1.96	1.87	1.00	0.99	-0.09	No
	8	95869	4	2.00	2.08	0.96	1.01	0.09	No
		68511	4	2.42	2.18	1.06	1.08	-0.23	No
	10	94887	4	2.16	2.05	0.86	0.93	-0.13	No
		94955	4	2.00	1.85	1.16	0.94	-0.13	No
Science	4	75427	4	1.23	1.24	0.91	0.85	0.02	No
	10	75652	4	0.92	0.69	0.93	0.76	-0.24	No

Appendix J—SCORE DISTRIBUTIONS

Table J-1. 2009–10 MontCAS: Performance Level Distributions by Subject and Grade

Subject	*Grade*	*Performance level*	*Percent in level:*		
			2009–10	*2008–09*	*2007–08*
Mathematics	3	4	28.93	27.80	25.22
		3	40.47	39.66	38.28
		2	16.11	16.03	19.54
		1	14.49	16.51	16.97
	4	4	31.91	28.04	27.01
		3	37.29	39.09	40.23
		2	17.02	18.05	18.92
		1	13.78	14.83	13.84
	5	4	32.57	32.97	26.08
		3	37.91	34.37	41.96
		2	16.49	18.51	18.83
		1	13.03	14.15	13.14
	6	4	32.41	33.09	25.68
		3	36.22	32.03	37.97
		2	17.72	18.66	20.03
		1	13.64	16.22	16.31
	7	4	35.93	31.27	29.83
		3	31.69	35.38	37.39
		2	16.38	18.76	20.00
		1	15.99	14.59	12.78
	8	4	26.91	27.44	25.98
		3	40.16	33.53	34.05
		2	22.94	23.81	25.91
		1	9.99	15.22	14.06
	10	4	21.66	18.93	18.14
		3	35.72	35.96	35.05
		2	31.24	35.27	35.44
		1	11.38	9.84	11.38
Reading	3	4	45.64	40.94	42.26
		3	39.40	43.78	41.93
		2	12.49	12.27	11.79
		1	2.47	3.01	4.02
	4	4	45.21	39.41	34.29
		3	38.18	42.21	45.28
		2	12.24	13.98	15.90
		1	4.36	4.40	4.53
	5	4	55.14	48.43	51.31
		3	31.34	35.94	30.83
		2	9.87	11.74	11.17
		1	3.65	3.89	6.69
	6	4	53.97	49.89	45.58
		3	32.65	35.21	38.41
		2	8.63	9.17	10.99
		1	4.75	5.72	5.02
	7	4	48.65	47.14	44.94
		3	35.73	36.42	38.77
		2	10.29	11.73	10.06
		1	5.33	4.71	6.24

continued

Subject	Grade	Performance level	Percent in level:		
			2009–10	2008–09	2007–08
Reading	8	4	54.49	48.57	45.88
		3	30.19	33.04	35.92
		2	9.51	11.26	10.17
		1	5.82	7.13	8.02
	10	4	43.91	43.70	35.35
		3	36.92	35.21	43.18
		2	9.63	11.66	13.26
		1	9.54	9.43	8.21
Science	4	4	15.80	11.61	14.34
		3	50.66	54.47	48.32
		2	27.72	27.01	30.29
		1	5.82	6.92	7.05
	8	4	18.42	14.16	12.40
		3	44.41	46.14	46.98
		2	27.66	28.43	29.06
		1	9.51	11.27	11.56
	10	4	17.78	18.25	16.21
		3	25.00	24.40	26.76
		2	33.59	33.69	34.35
		1	23.62	23.66	22.68

Figure J-1. 2009–10 MontCAS: Scaled Score Percentages – Mathematics Grade 3

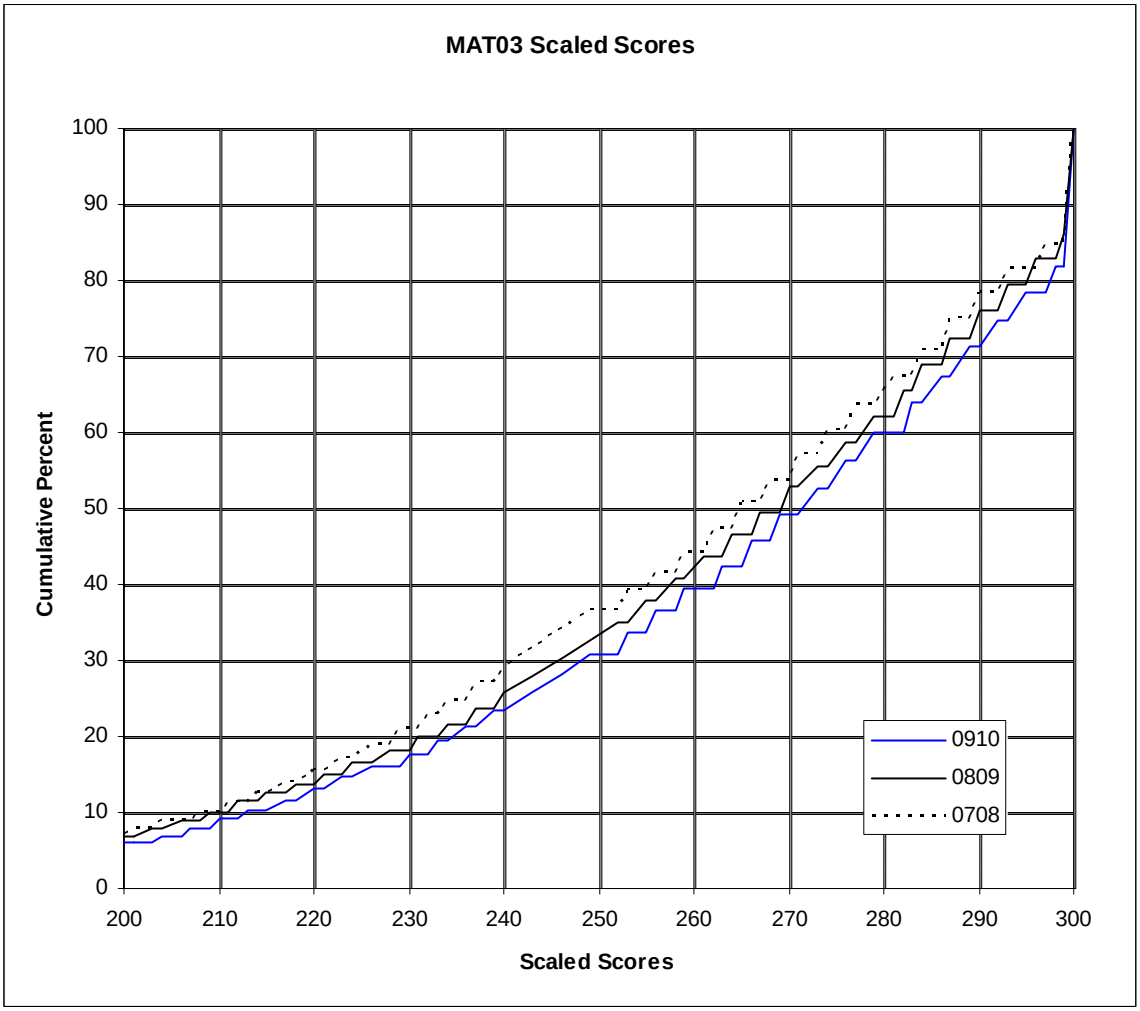


Figure J-2. 2009–10 MontCAS: Scaled Score Percentages – Mathematics Grade 4

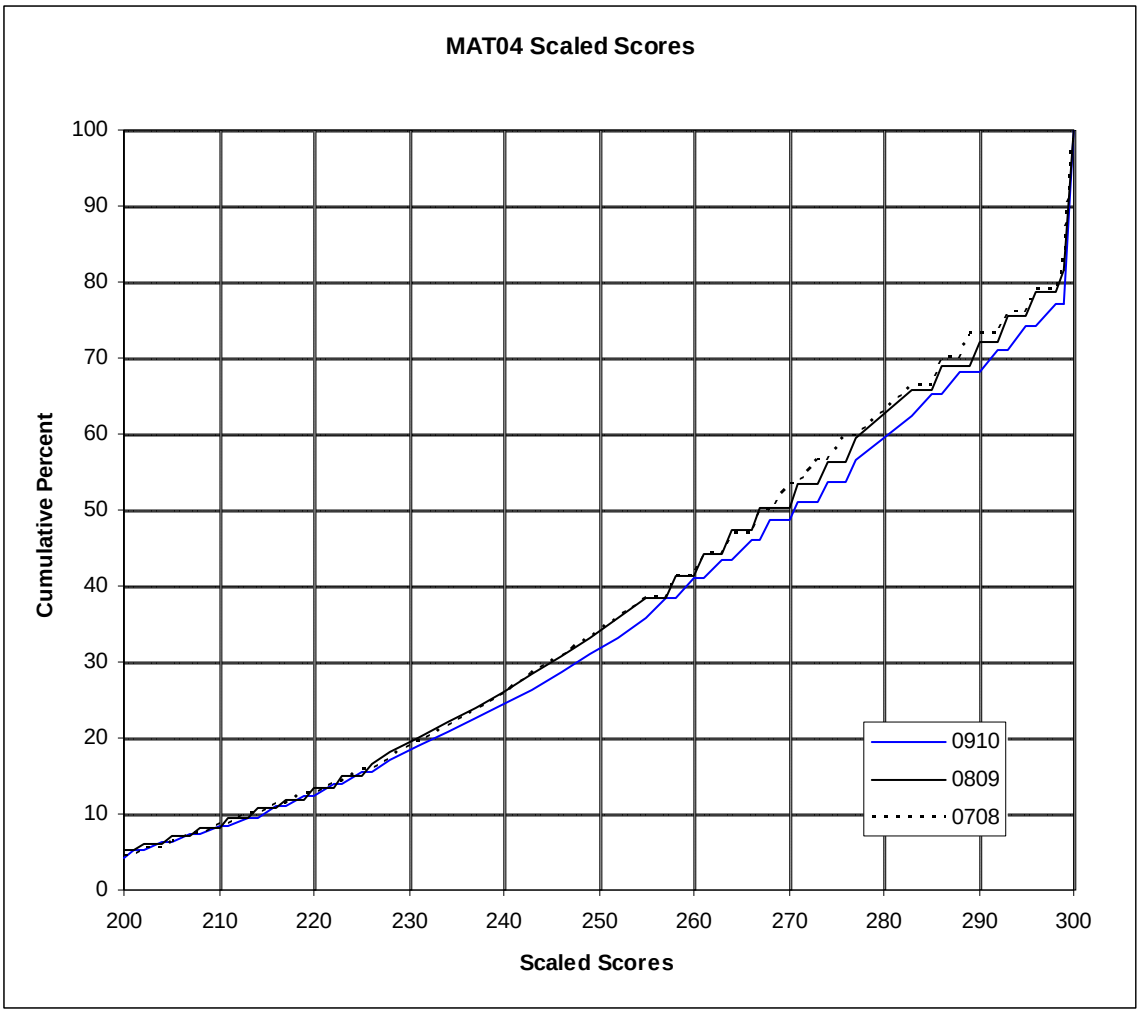


Figure J-3. 2009–10 MontCAS: Scaled Score Percentages – Mathematics Grade 5

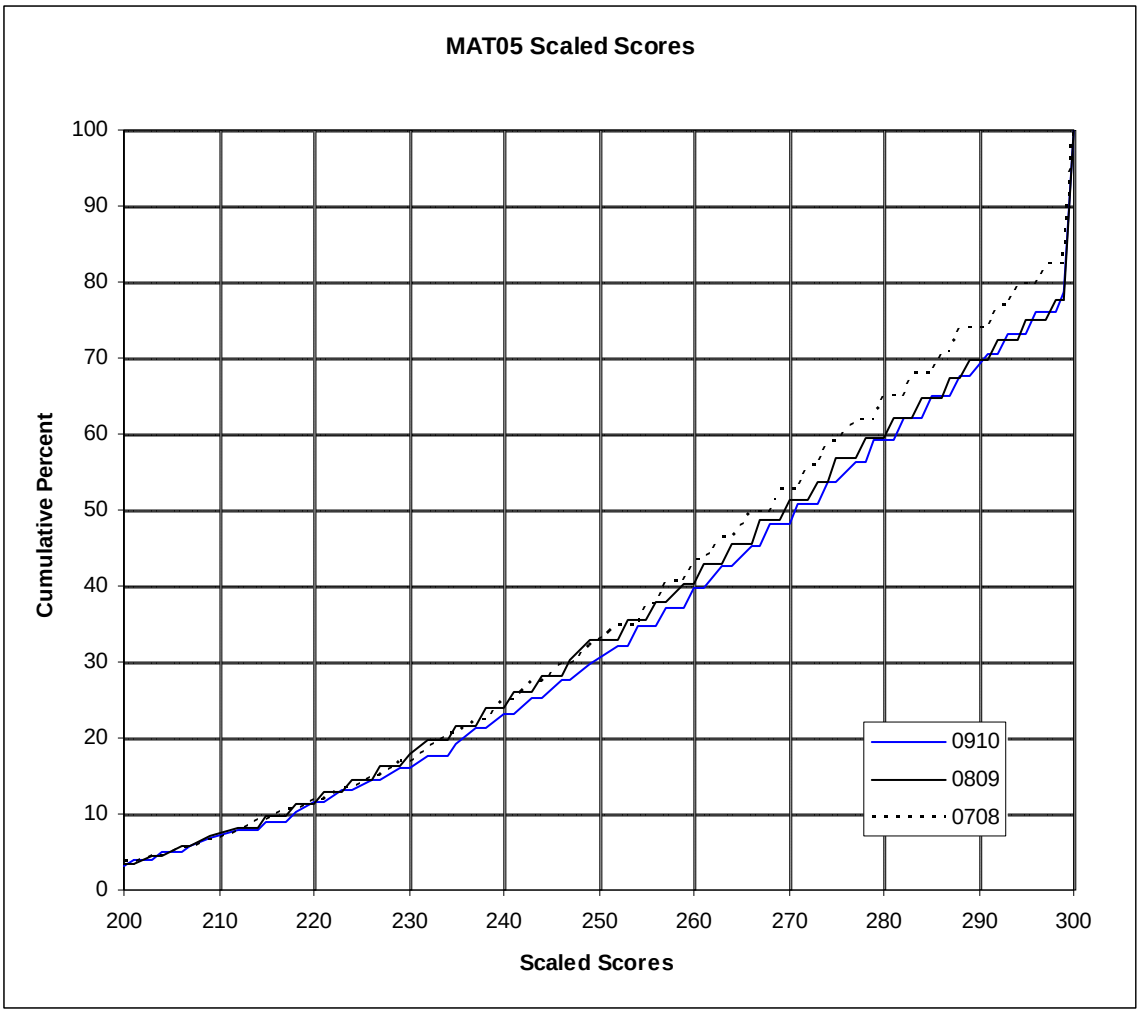


Figure J-4. 2009–10 MontCAS: Scaled Score Percentages – Mathematics Grade 6

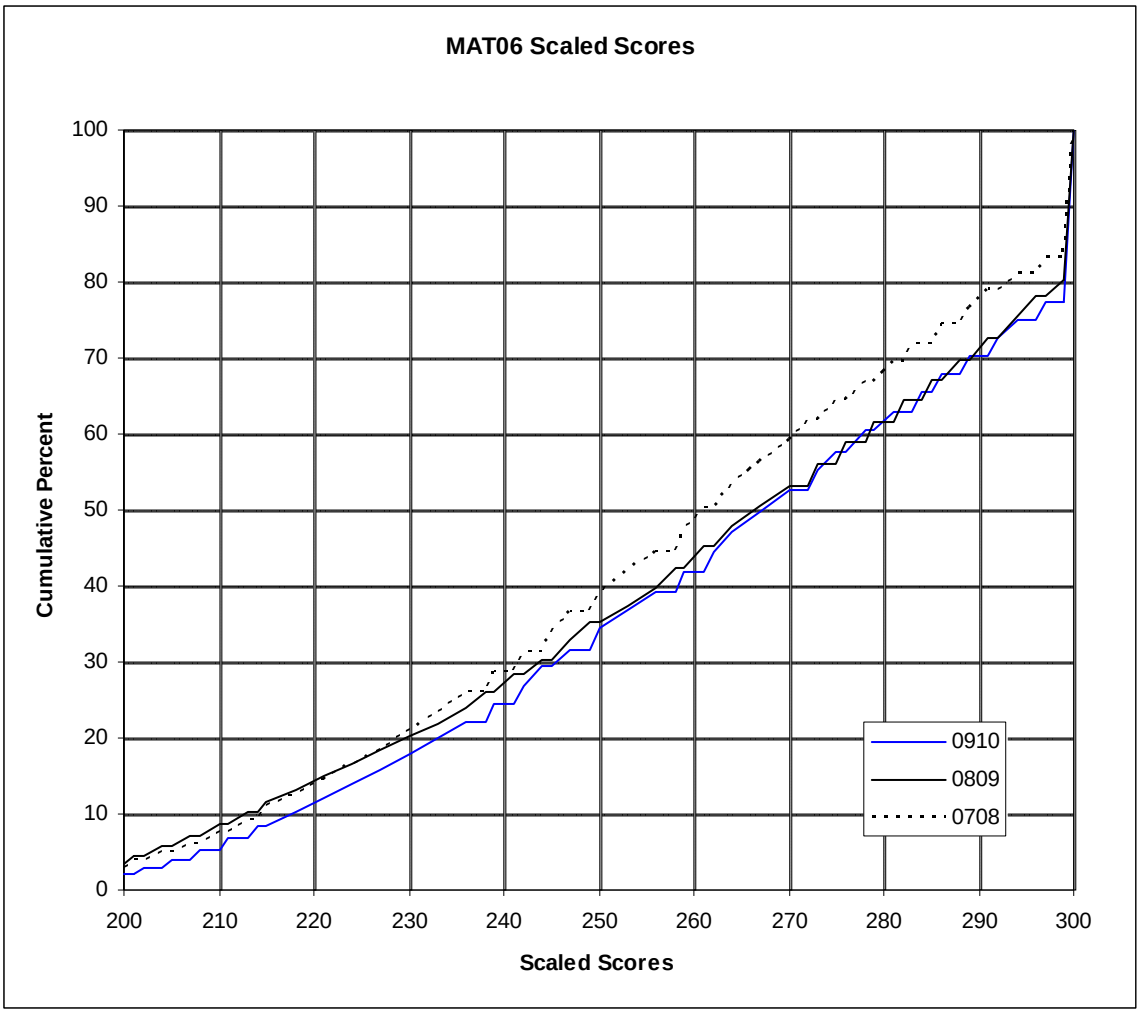


Figure J-5. 2009–10 MontCAS: Scaled Score Percentages – Mathematics Grade 7

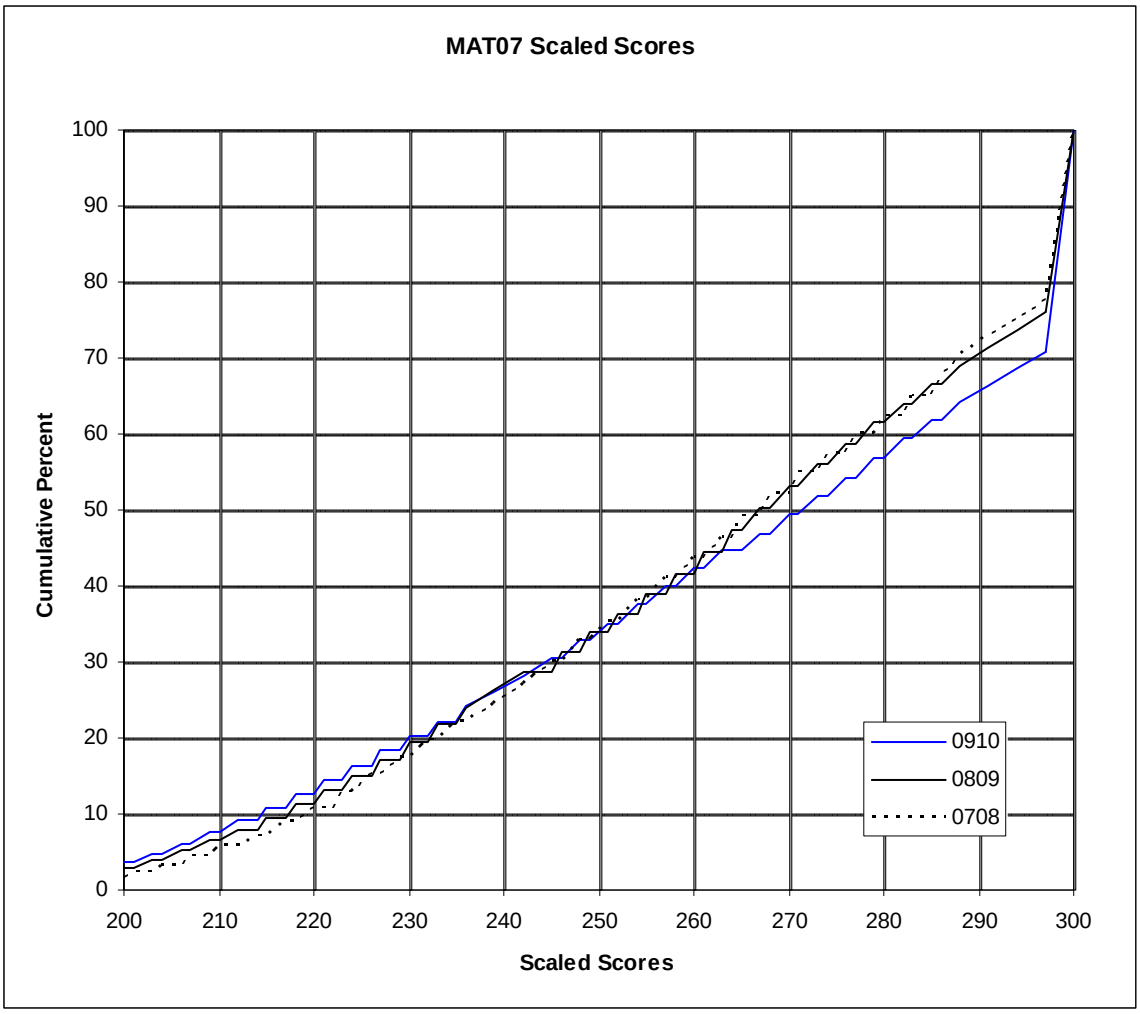


Figure J-6. 2009–10 MontCAS: Scaled Score Percentages – Mathematics Grade 8

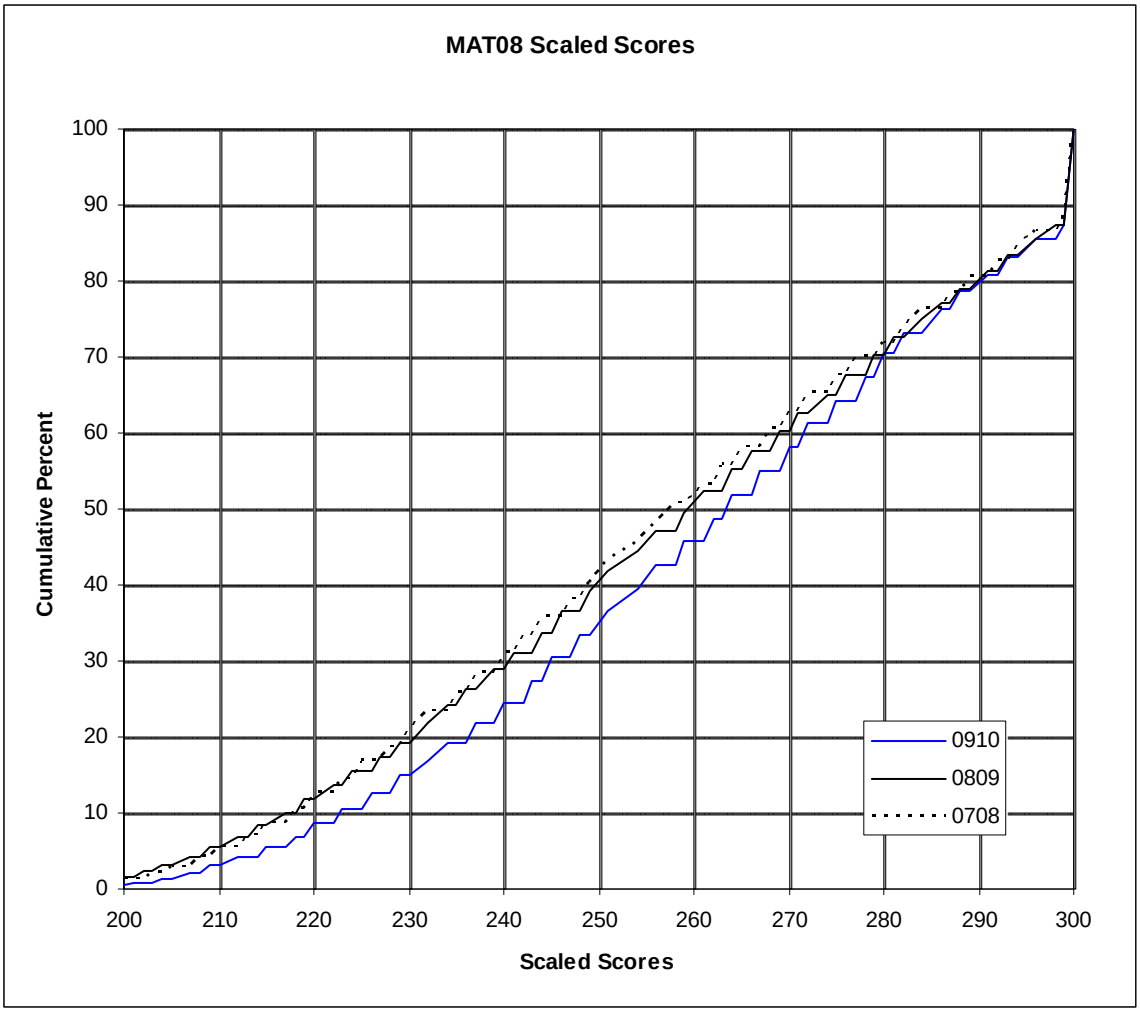


Figure J-7. 2009–10 MontCAS: Scaled Score Percentages – Mathematics Grade 10

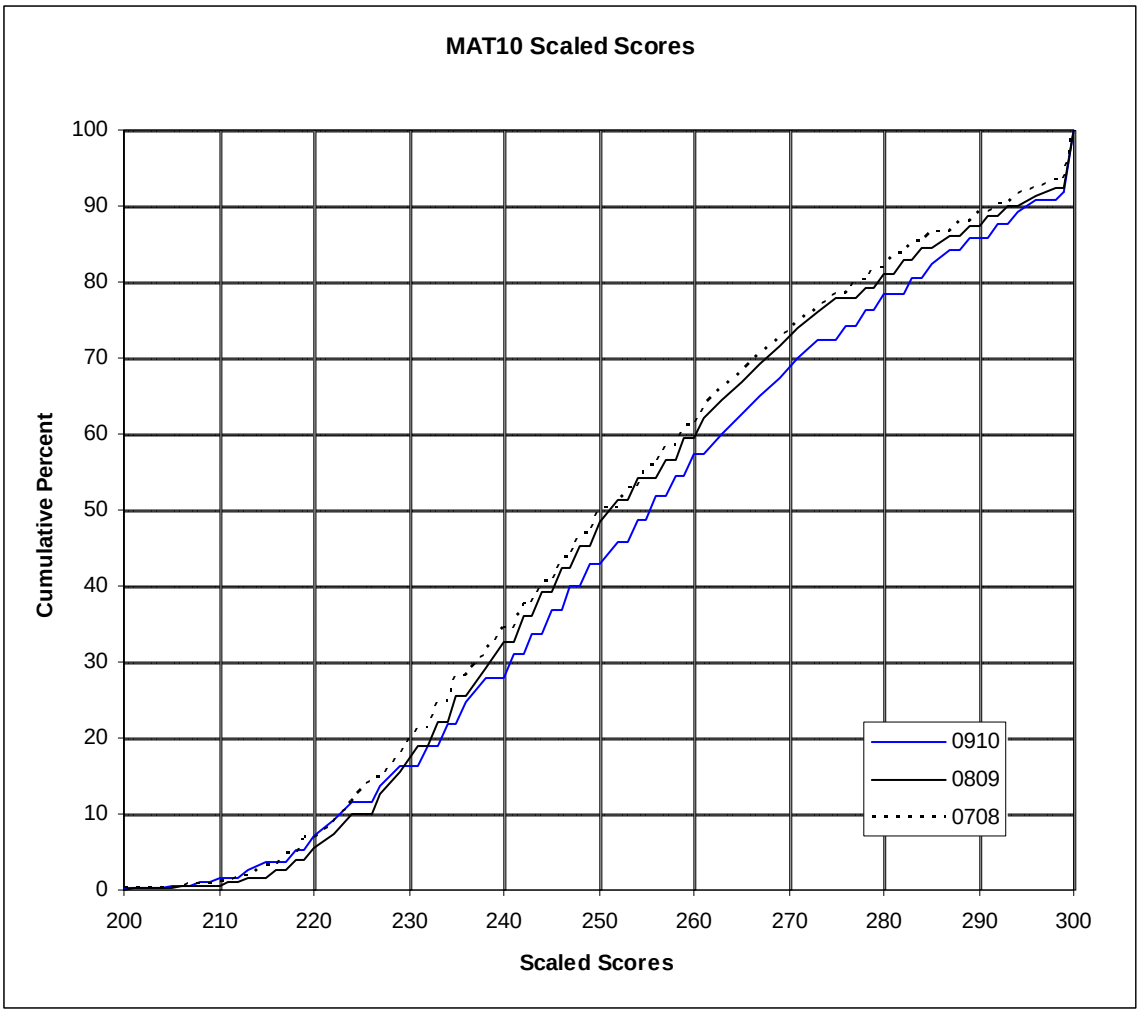


Figure J-8. 2009–10 MontCAS: Scaled Score Percentages – Reading Grade 3

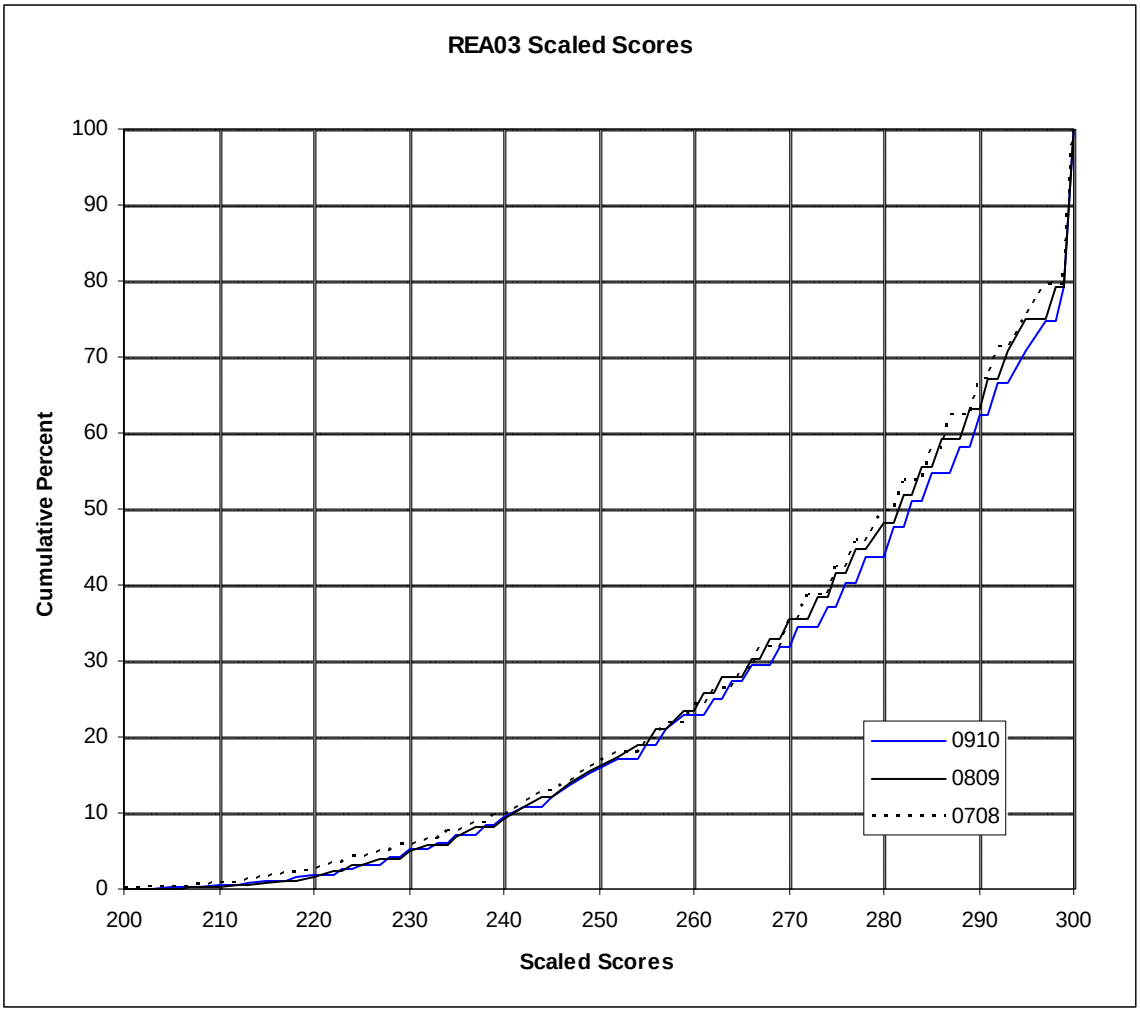


Figure J-9. 2009–10 MontCAS: Scaled Score Percentages – Reading Grade 4

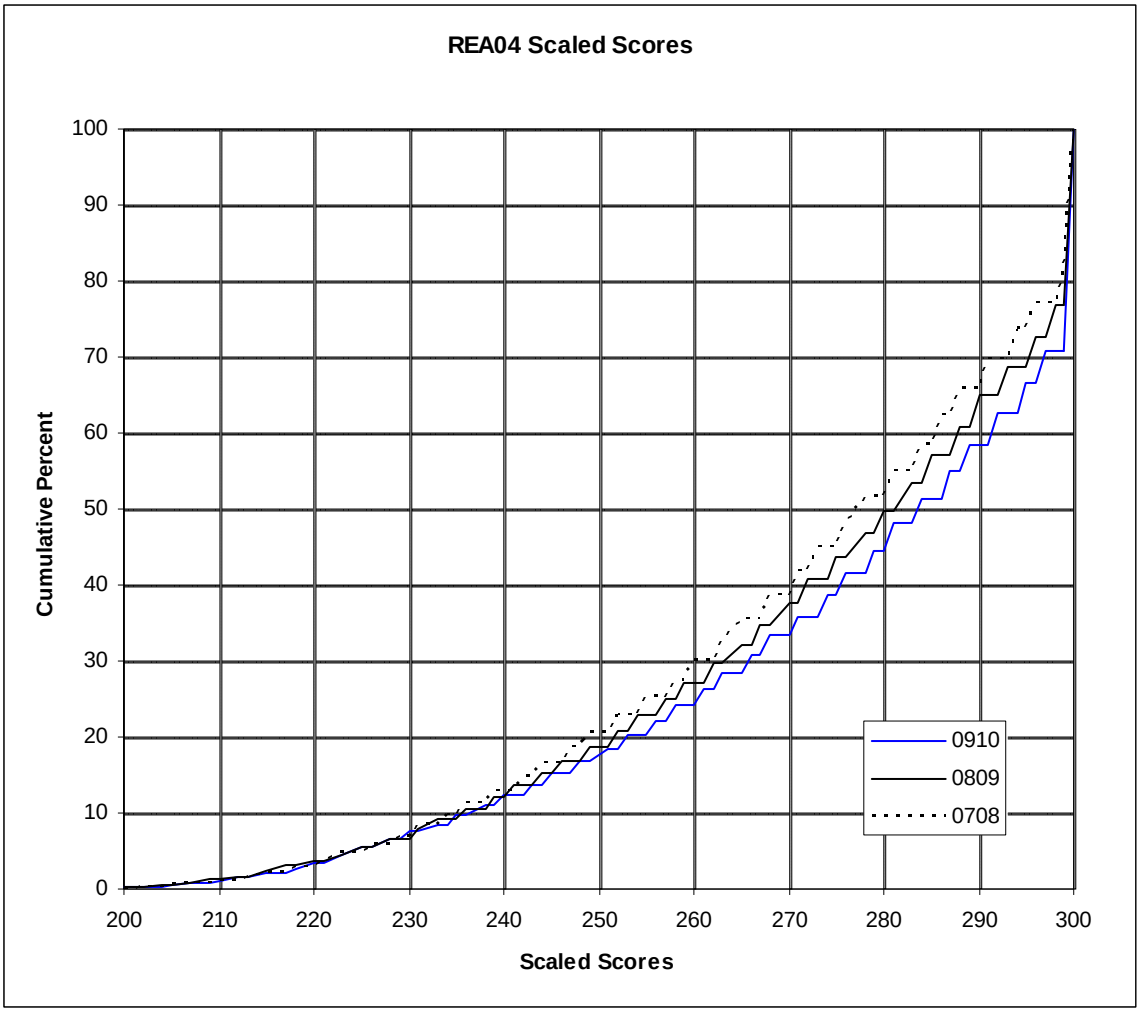


Figure J-10. 2009–10 MontCAS: Scaled Score Percentages – Reading Grade 5

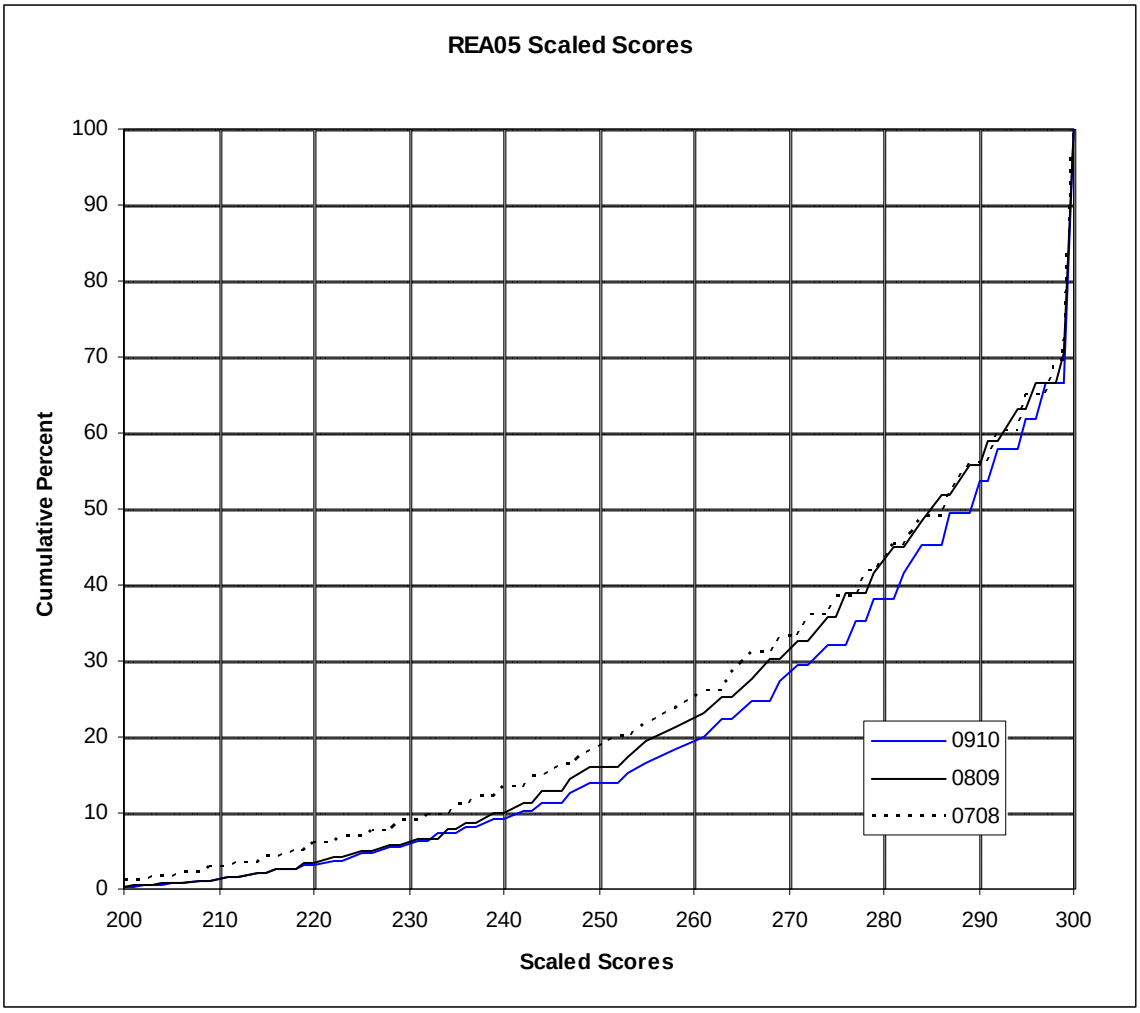


Figure J-11. 2009–10 MontCAS: Scaled Score Percentages – Reading Grade 6

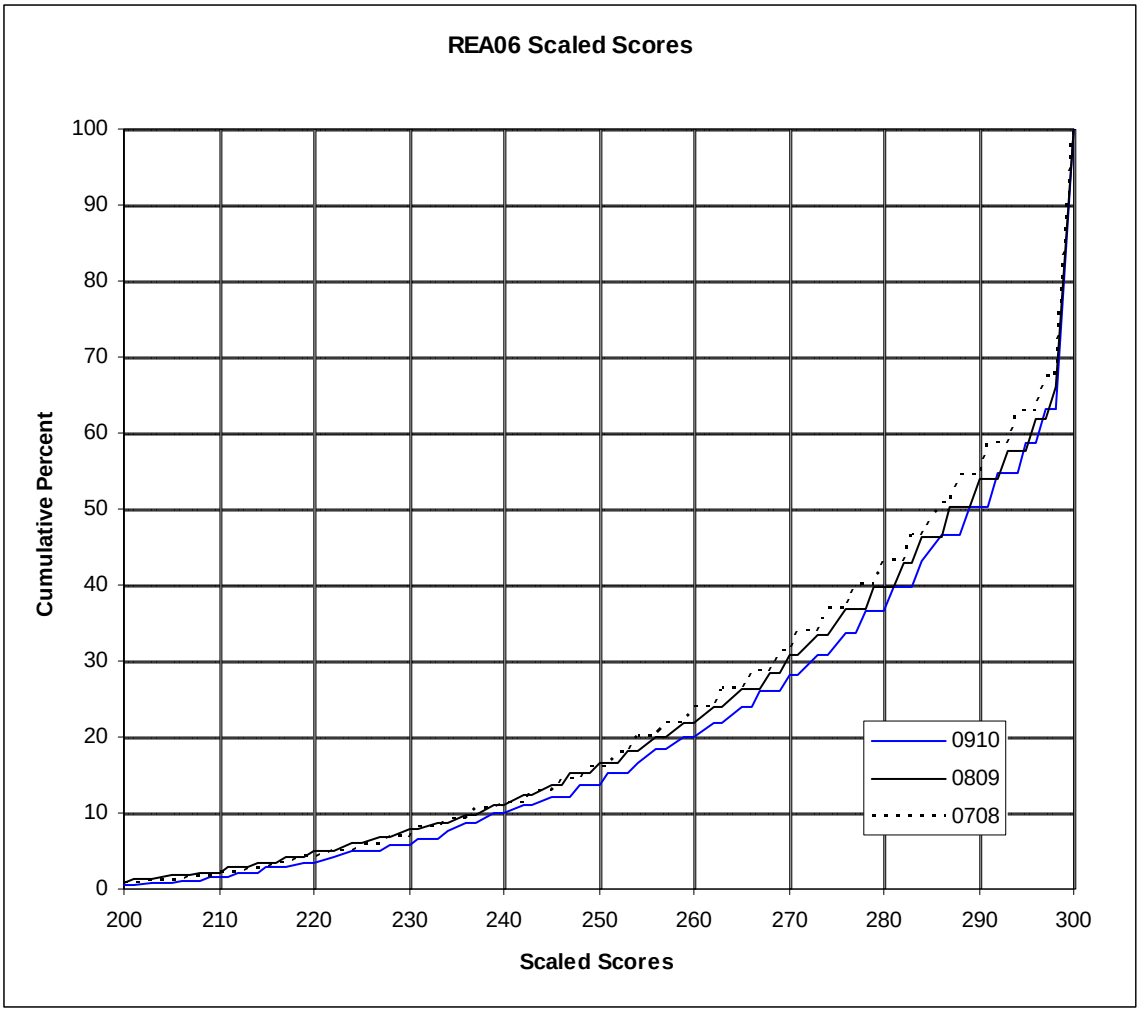


Figure J-12. 2009–10 MontCAS: Scaled Score Percentages – Reading Grade 7

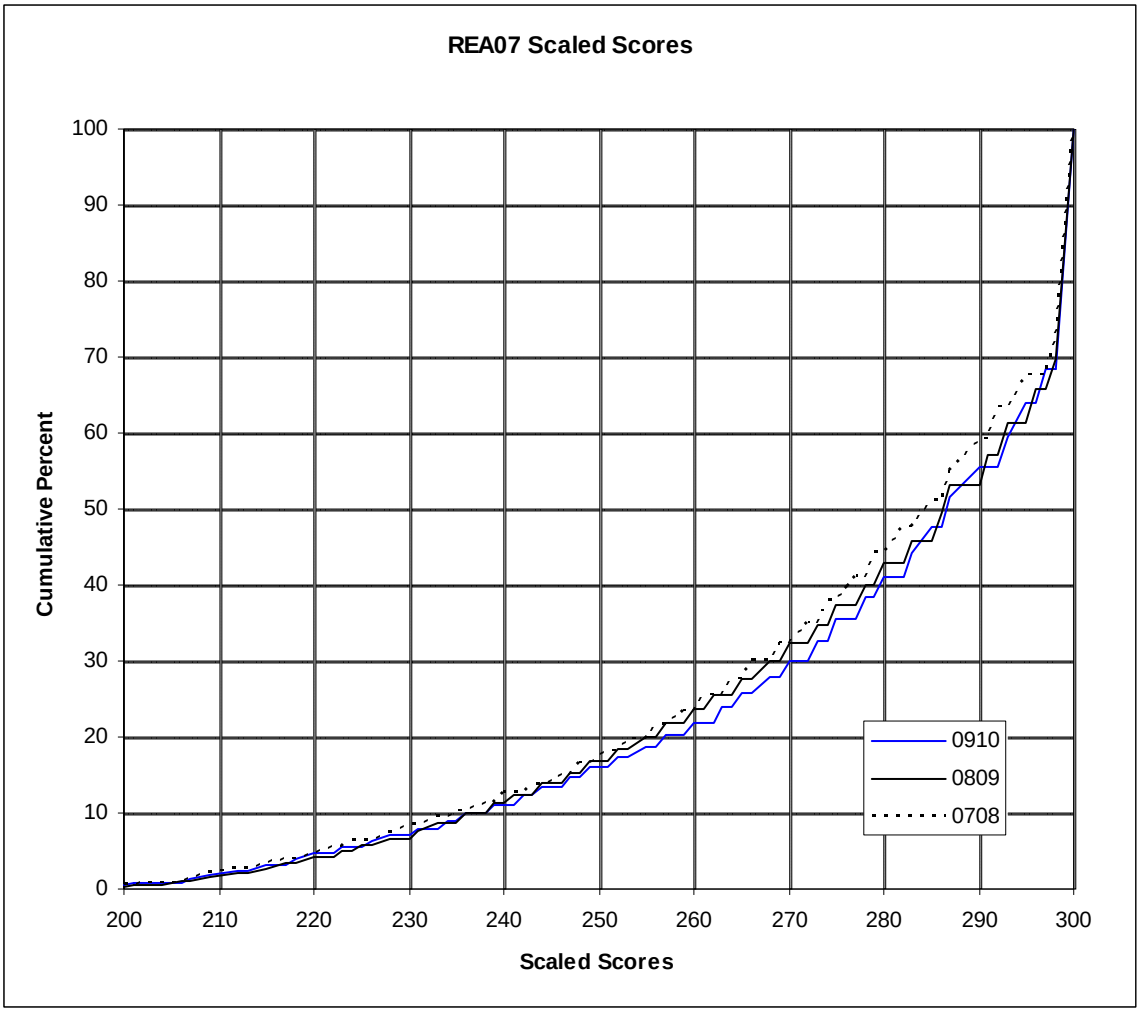


Figure J-13. 2009–10 MontCAS: Scaled Score Percentages – Reading Grade 8

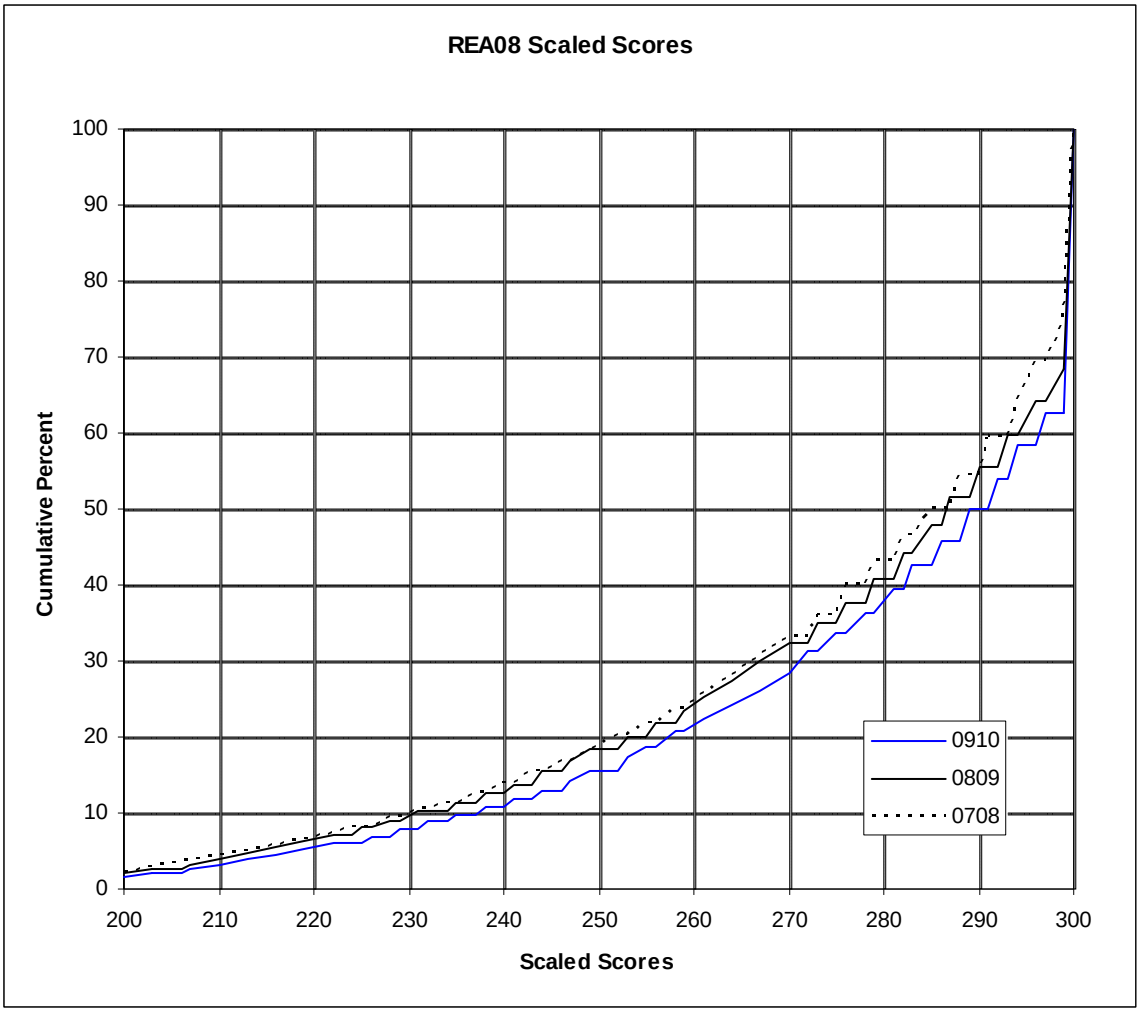


Figure J-14. 2009–10 MontCAS: Scaled Score Percentages – Reading Grade 10

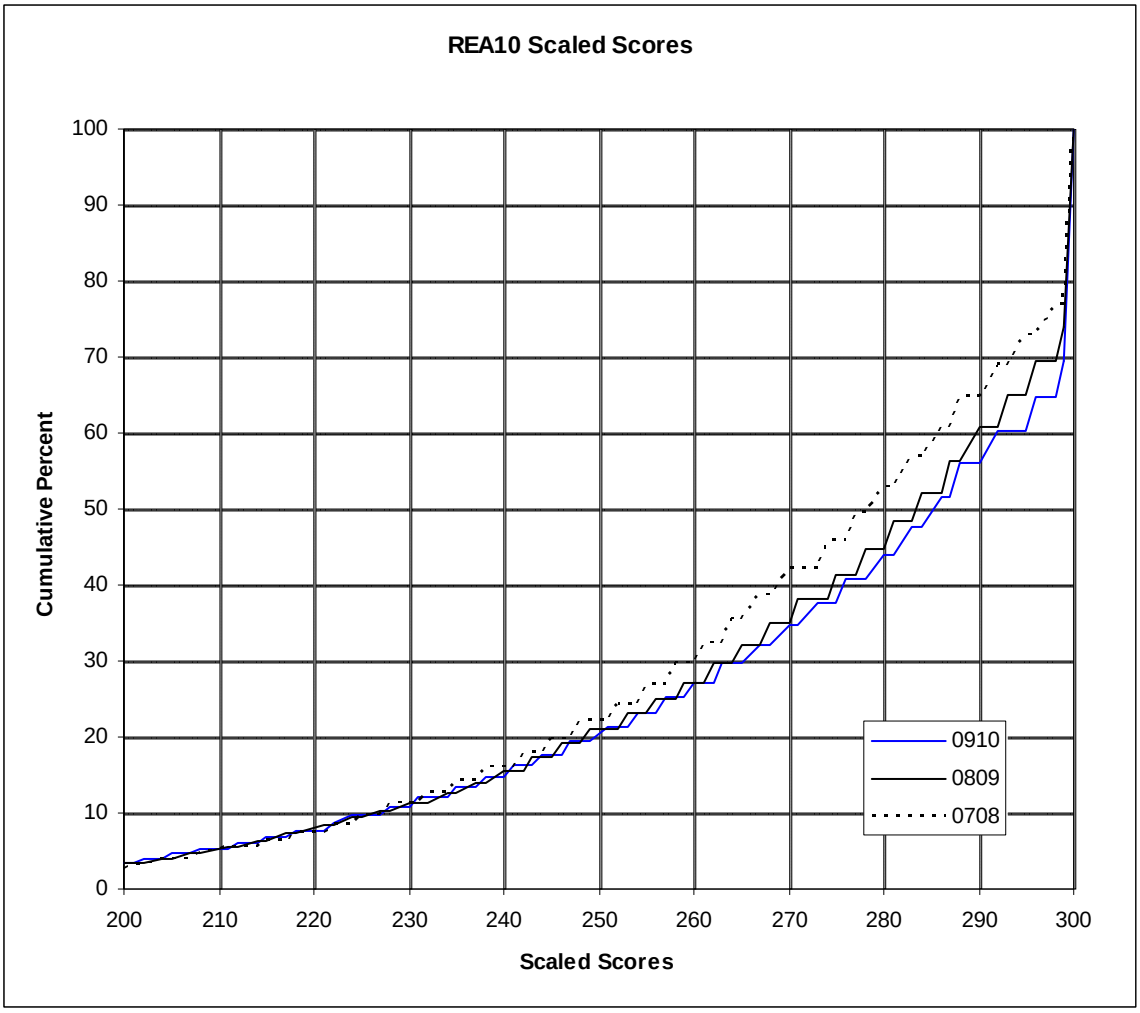


Figure J-15. 2009–10 MontCAS: Scaled Score Percentages – Science Grade 4

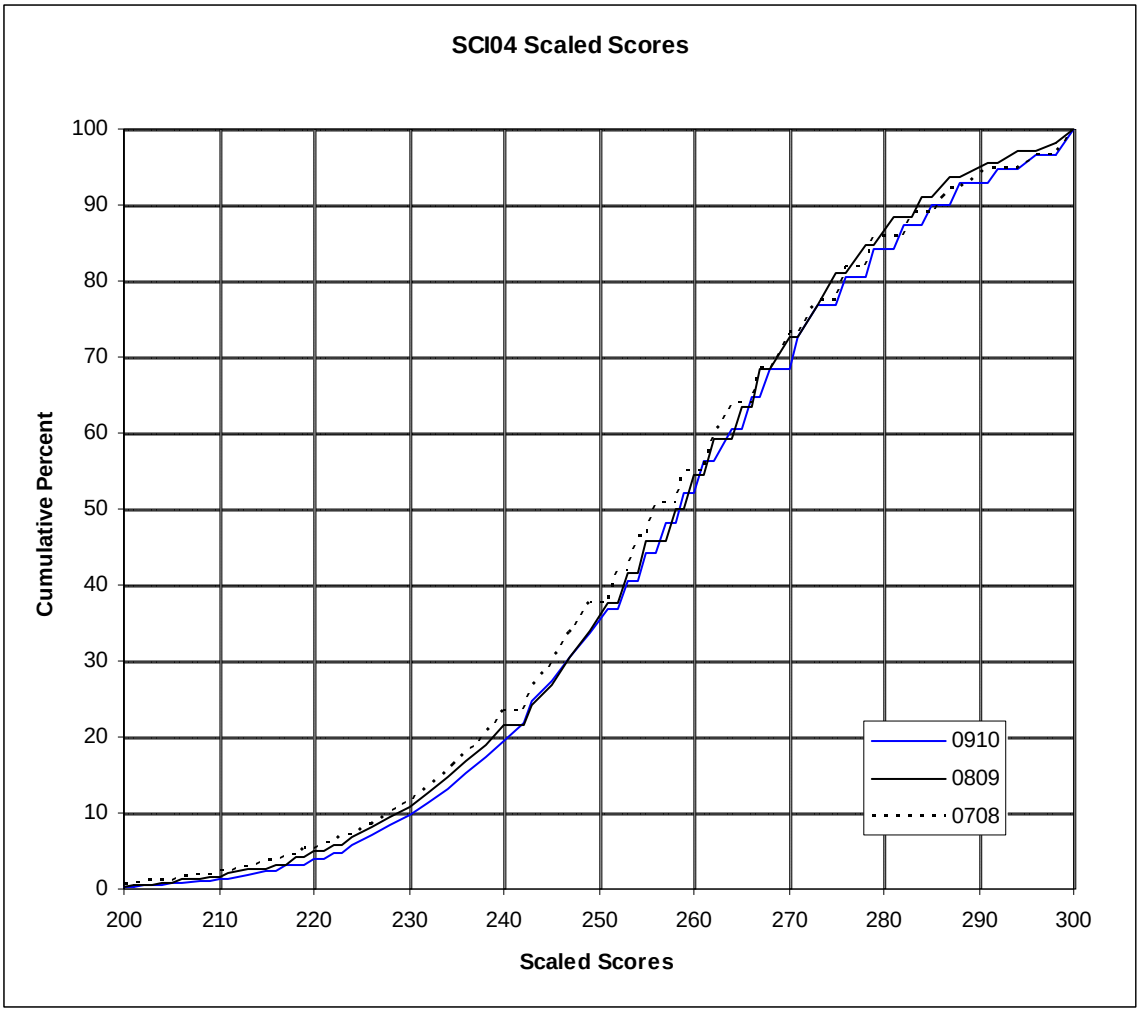


Figure J-16. 2009–10 MontCAS: Scaled Score Percentages – Science Grade 8

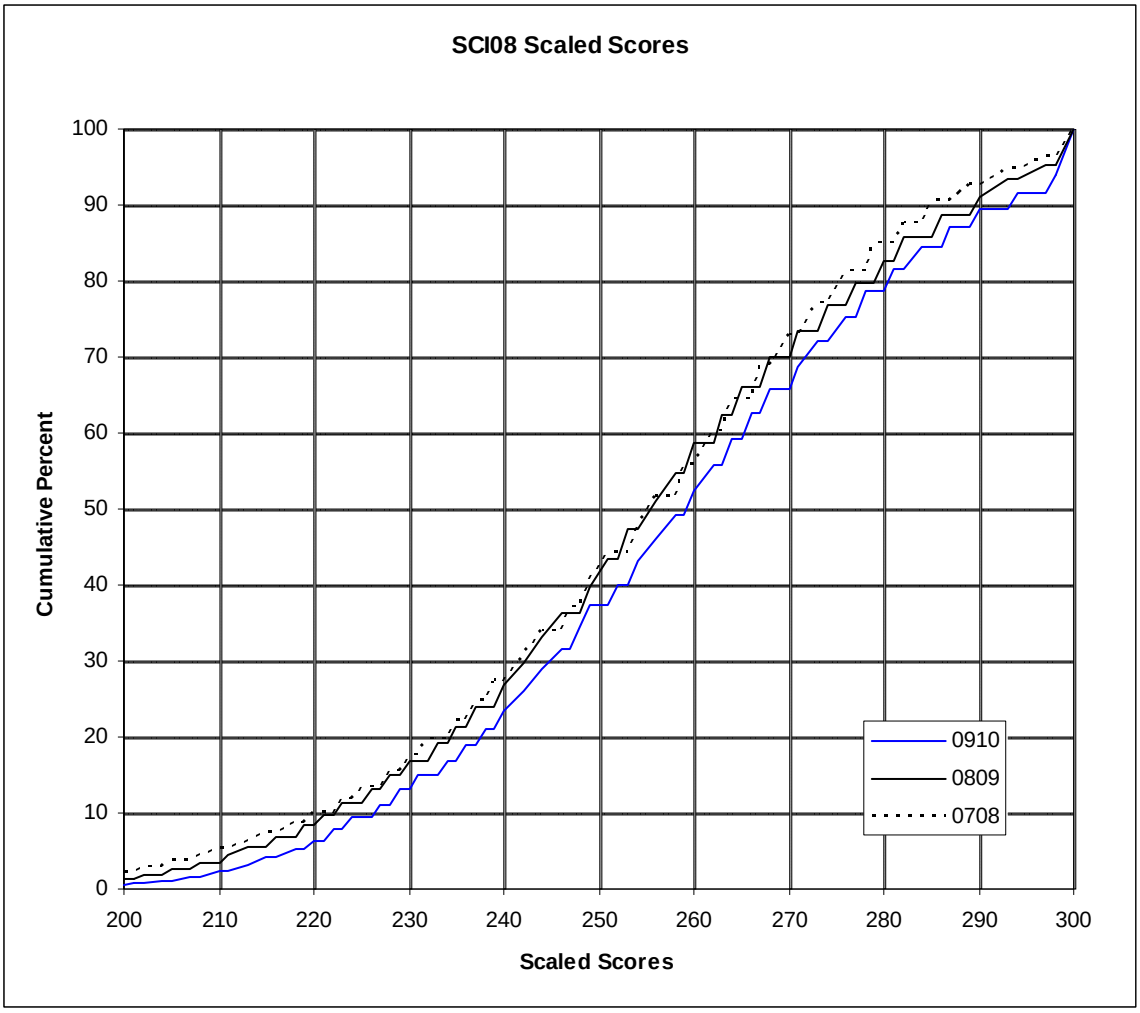
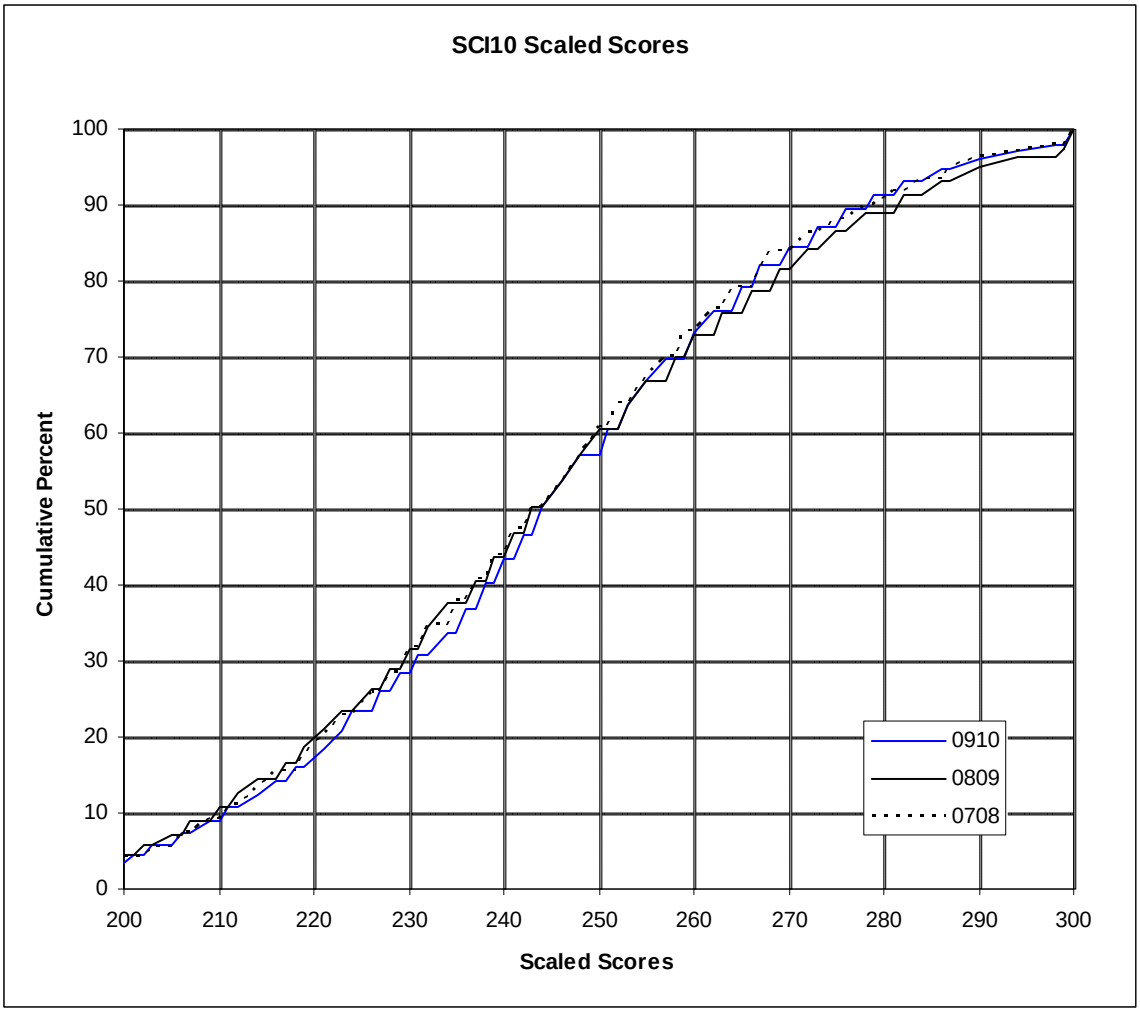


Figure J-17. 2009–10 MontCAS: Scaled Score Percentages – Science Grade 10



Appendix K—Raw to Scaled Score Look-up Tables

Table K-1. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 3

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	200
14	200	200
15	200	200
16	200	200
17	200	200
18	200	200
19	200	200
20	200	200
21	200	200
22	200	200
23	200	200
24	200	200
25	203	200
26	206	204
27	209	207
28	212	210
29	215	213
30	218	217
31	221	220
32	224	223
33	228	226
34	231	230
35	234	233
36	237	236
37	240	239
38	243	243
39	246	246
40	249	249
41	252	253
42	255	256
43	258	259
44	261	263
45	264	266
46	267	269
47	270	273
48	273	276
49	276	279
50	279	283
51	282	286
52	284	289
53	287	292
54	290	295
55	293	298

continued

Raw score	*2008–09*	*2009–10*
56	296	300
57	299	300
58	300	300
59	300	300
60	300	300
61	300	300
62	300	300
63	300	300
64	300	300
65	300	300
66	300	300

Table K-2. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 4

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	200
14	200	200
15	200	200
16	200	200
17	200	200
18	200	200
19	200	200
20	200	201
21	202	204
22	205	207
23	208	210
24	211	213
25	214	216
26	217	219
27	220	222
28	223	225
29	226	228
30	228	231
31	231	234
32	234	237
33	237	240
34	240	243
35	243	246
36	246	249
37	249	252
38	252	255
39	255	257
40	258	260
41	261	263
42	264	266

continued

Raw score	*2008–09*	*2009–10*
43	267	268
44	271	271
45	274	274
46	277	277
47	280	280
48	283	283
49	286	285
50	290	288
51	293	292
52	296	295
53	299	298
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300
62	300	300
63	300	300
64	300	300
65	300	300
66	300	300

Table K-3. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 5

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	200
14	200	200
15	200	200
16	200	200
17	203	201
18	206	204
19	209	207
20	212	209
21	215	212
22	218	215
23	221	218
24	224	220
25	227	223
26	230	226
27	232	229
28	235	232
29	238	235

continued

Raw score	*2008–09*	*2009–10*
30	241	237
31	244	240
32	247	243
33	249	246
34	253	249
35	256	252
36	259	254
37	261	257
38	264	260
39	267	263
40	270	266
41	273	268
42	275	271
43	278	274
44	241	277
45	284	279
46	287	282
47	289	285
48	292	288
49	295	291
50	298	293
51	300	296
52	300	299
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300
62	300	300
63	300	300
64	300	300
65	300	300
66	300	300

Table K-4. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 6

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	200
14	200	200
15	201	202
16	204	205

continued

Raw score	*2008–09*	*2009–10*
17	207	208
18	210	211
19	213	214
20	215	218
21	218	221
22	221	224
23	224	227
24	227	230
25	230	233
26	233	236
27	236	239
28	238	242
29	241	244
30	244	247
31	247	250
32	249	253
33	253	256
34	256	259
35	258	262
36	261	264
37	264	267
38	267	270
39	270	273
40	273	275
41	276	278
42	279	281
43	282	284
44	285	286
45	288	289
46	291	292
47	294	294
48	296	297
49	299	300
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300
62	300	300
63	300	300
64	300	300
65	300	300
66	300	300

Table K-5. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 7

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200

continued

Raw score	*2008–09*	*2009–10*
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	200
14	203	200
15	206	203
16	209	206
17	212	209
18	215	212
19	218	215
20	221	218
21	224	221
22	228	224
23	231	227
24	234	230
25	237	233
26	240	236
27	243	239
28	246	242
29	249	245
30	253	248
31	256	251
32	259	254
33	262	257
34	265	260
35	268	263
36	272	267
37	275	270
38	278	273
39	281	276
40	284	279
41	287	282
42	290	285
43	293	288
44	296	291
45	298	294
46	300	297
47	300	300
48	300	300
49	300	300
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300

continued

Raw score	*2008–09*	*2009–10*
62	300	300
63	300	300
64	300	300
65	300	300
66	300	300

Table K-6. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 8

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	201
12	200	204
13	202	207
14	204	209
15	207	212
16	209	215
17	212	218
18	214	220
19	217	223
20	219	226
21	222	229
22	224	232
23	227	234
24	229	237
25	232	240
26	234	243
27	236	245
28	239	248
29	241	251
30	244	254
31	246	256
32	249	259
33	251	262
34	254	264
35	256	267
36	259	270
37	261	272
38	264	275
39	266	278
40	269	280
41	271	282
42	274	286
43	276	288
44	279	291
45	281	293
46	284	296
47	286	299
48	288	300

continued

Raw score	*2008–09*	*2009–10*
49	291	300
50	293	300
51	296	300
52	298	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300
62	300	n/a
63	300	n/a
64	300	n/a
65	300	n/a
66	300	n/a

Table K-7. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Mathematics Grade 10

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	201	203
9	204	205
10	206	208
11	209	210
12	211	213
13	213	215
14	216	218
15	218	220
16	220	222
17	222	224
18	224	227
19	227	229
20	229	232
21	231	234
22	233	236
23	235	238
24	238	241
25	240	243
26	242	245
27	244	247
28	246	249
29	248	252
30	250	254
31	252	256
32	254	258
33	257	260
34	259	263
35	261	265

continued

Raw score	*2008–09*	*2009–10*
36	263	267
37	265	269
38	267	271
39	269	273
40	271	276
41	273	278
42	275	280
43	278	283
44	280	285
45	282	287
46	284	289
47	287	292
48	289	294
49	291	296
50	293	299
51	296	300
52	298	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300
62	300	300
63	300	300
64	300	300
65	300	300
66	300	300

Table K-8. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 3

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	202	200
8	205	203
9	207	205
10	210	208
11	212	210
12	215	213
13	217	215
14	220	218
15	222	220
16	224	223
17	227	225
18	230	228
19	232	230
20	235	233
21	237	235
22	240	238

continued

Raw score	*2008–09*	*2009–10*
23	242	240
24	244	242
25	247	245
26	249	247
27	252	249
28	254	252
29	256	255
30	259	257
31	261	259
32	263	262
33	266	264
34	268	266
35	270	269
36	273	271
37	275	274
38	277	276
39	280	278
40	282	281
41	284	283
42	286	285
43	289	288
44	291	290
45	293	292
46	295	295
47	298	297
48	300	299
49	300	300
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300

Table K-9. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 4

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	201	202
11	204	205
12	207	207
13	209	210
14	212	212
15	215	215

continued

Raw score	*2008–09*	*2009–10*
16	217	218
17	220	220
18	223	223
19	225	225
20	228	228
21	231	230
22	233	233
23	236	235
24	239	238
25	241	240
26	244	243
27	246	245
28	249	248
29	252	251
30	254	253
31	257	256
32	259	258
33	262	261
34	265	263
35	267	266
36	270	268
37	272	271
38	275	274
39	278	276
40	280	279
41	283	281
42	285	284
43	288	287
44	290	289
45	293	292
46	296	295
47	298	297
48	300	300
49	300	300
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300

Table K-10. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 5

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200

continued

Raw score	*2008–09*	*2009–10*
9	200	200
10	200	200
11	201	200
12	204	202
13	208	205
14	211	208
15	214	211
16	216	214
17	219	216
18	222	219
19	225	222
20	228	225
21	231	228
22	234	231
23	236	233
24	239	236
25	242	239
26	244	242
27	247	244
28	249	247
29	253	249
30	255	253
31	258	255
32	261	258
33	263	261
34	266	263
35	268	266
36	271	269
37	274	271
38	276	274
39	279	277
40	281	279
41	284	282
42	286	284
43	289	287
44	291	290
45	294	292
46	296	295
47	299	297
48	300	300
49	300	300
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300

Table K-11. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 6

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	201	203
14	205	206
15	208	209
16	211	212
17	214	215
18	217	219
19	220	222
20	224	224
21	227	228
22	230	231
23	233	234
24	236	236
25	239	239
26	242	242
27	245	245
28	247	248
29	250	251
30	253	254
31	256	256
32	259	259
33	262	262
34	265	265
35	268	267
36	270	270
37	273	273
38	276	276
39	279	278
40	282	281
41	284	284
42	287	286
43	290	289
44	293	292
45	296	295
46	298	297
47	300	300
48	300	300
49	300	300
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300

continued

Raw score	*2008–09*	*2009–10*
57	300	300
58	300	300
59	300	300
60	300	300

Table K-12. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 7

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	201	201
12	204	204
13	206	207
14	209	209
15	212	212
16	215	215
17	217	218
18	220	220
19	223	223
20	225	226
21	228	228
22	231	231
23	233	234
24	236	236
25	239	239
26	241	242
27	244	244
28	247	247
29	249	249
30	252	252
31	255	255
32	257	257
33	260	260
34	262	263
35	265	265
36	268	268
37	270	270
38	273	273
39	275	275
40	278	278
41	280	280
42	283	283
43	286	285
44	287	287
45	291	290
46	293	293
47	296	295
48	298	297
49	300	300

continued

Raw score	*2008–09*	*2009–10*
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300

Table K-13. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 8

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	200
14	200	200
15	200	200
16	200	203
17	203	207
18	207	210
19	210	213
20	213	216
21	216	219
22	219	222
23	222	226
24	225	229
25	228	232
26	231	235
27	235	238
28	238	241
29	241	244
30	244	247
31	247	249
32	249	253
33	253	255
34	256	258
35	259	261
36	261	264
37	264	267
38	267	270
39	270	272
40	273	275
41	276	278
42	279	281

continued

Raw score	*2008–09*	*2009–10*
43	282	283
44	285	286
45	287	289
46	290	292
47	293	294
48	296	297
49	299	300
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300

Table K-14. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Reading Grade 10

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	200
14	200	200
15	200	200
16	200	200
17	200	200
18	200	200
19	204	202
20	207	205
21	211	208
22	214	212
23	217	215
24	221	218
25	224	222
26	227	224
27	230	228
28	234	231
29	237	235
30	240	238
31	243	241
32	246	244
33	249	247
34	253	251
35	256	254

continued

Raw score	*2008–09*	*2009–10*
36	259	257
37	262	260
38	265	263
39	268	267
40	271	270
41	275	273
42	278	276
43	281	280
44	284	283
45	287	286
46	290	288
47	293	292
48	296	296
49	299	299
50	300	300
51	300	300
52	300	300
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300

Table K-15. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Science Grade 4

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200
13	200	202
14	200	205
15	201	208
16	204	210
17	206	213
18	209	215
19	211	217
20	213	220
21	216	222
22	218	224
23	220	226
24	222	228
25	224	230
26	226	232
27	228	234
28	230	236

continued

Raw score	*2008–09*	*2009–10*
29	232	238
30	234	240
31	236	242
32	238	243
33	240	245
34	243	247
35	245	249
36	247	251
37	249	253
38	251	255
39	253	257
40	255	259
41	258	261
42	260	264
43	262	266
44	265	268
45	267	271
46	270	273
47	273	276
48	275	279
49	278	282
50	281	285
51	284	288
52	287	292
53	291	296
54	294	300
55	298	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300

Table K-16. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Science Grade 8

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	201
13	200	204
14	200	207
15	202	210
16	205	213
17	208	215
18	211	218
19	213	220
20	216	222

continued

Raw score	*2008–09*	*2009–10*
21	219	224
22	221	227
23	223	229
24	226	231
25	228	234
26	230	236
27	233	238
28	235	240
29	237	242
30	240	244
31	242	246
32	244	248
33	246	249
34	249	252
35	251	254
36	253	256
37	256	258
38	258	260
39	260	262
40	263	264
41	265	266
42	268	268
43	271	271
44	274	273
45	277	276
46	280	278
47	282	281
48	286	284
49	290	287
50	293	290
51	297	294
52	300	298
53	300	300
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300

Table K-17. 2009–10 MontCAS: Raw to Scaled Score Look-up Table – Science Grade 10

Raw score	*2008–09*	*2009–10*
0	200	200
1	200	200
2	200	200
3	200	200
4	200	200
5	200	200
6	200	200
7	200	200
8	200	200
9	200	200
10	200	200
11	200	200
12	200	200

continued

Raw score	*2008–09*	*2009–10*
13	200	200
14	200	200
15	200	201
16	200	203
17	202	206
18	205	209
19	207	211
20	210	214
21	212	216
22	214	218
23	217	221
24	219	223
25	221	224
26	223	227
27	226	229
28	228	231
29	230	234
30	232	236
31	234	238
32	237	240
33	239	242
34	241	244
35	243	246
36	246	248
37	248	251
38	250	253
39	253	255
40	255	257
41	258	260
42	260	262
43	263	265
44	266	267
45	269	270
46	272	273
47	275	276
48	278	279
49	282	282
50	286	286
51	290	290
52	294	294
53	299	298
54	300	300
55	300	300
56	300	300
57	300	300
58	300	300
59	300	300
60	300	300
61	300	300

APPENDIX L—CLASSICAL RELIABILITY

Table L-1. 2009–10 MontCAS: Subgroup Reliabilities – Mathematics

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
3	Special Education	1197	66	36.37	12.60	0.92	3.49
	Title 1	4096	66	42.75	11.38	0.91	3.37
	Low Income	4906	66	41.92	11.18	0.91	3.39
	American Indian	1317	66	37.57	11.52	0.91	3.49
	Asian	94	66	48.83	11.29	0.93	3.06
	Hispanic	318	66	42.87	10.74	0.90	3.39
	Black or African American	150	66	41.39	11.39	0.91	3.38
	White, Non-Hispanic	8625	66	46.39	10.33	0.90	3.24
	Native Hawaiian/Other Pacific Islander	40	66	44.28	11.36	0.92	3.27
	Female	5166	66	45.10	10.79	0.91	3.28
	Male	5378	66	45.16	11.08	0.91	3.28
	Limited English Proficient	291	66	32.73	10.84	0.89	3.57
	Migrant	25	66	39.04	12.50	0.92	3.48
	Plan 504	39	66	46.28	10.61	0.90	3.29
	All Students	10544	66	45.13	10.94	0.91	3.28
4	Special Education	1227	66	31.90	12.93	0.91	3.77
	Title 1	4095	66	39.24	12.93	0.92	3.77
	Low Income	4689	66	38.98	12.74	0.91	3.77
	American Indian	1276	66	34.64	12.42	0.91	3.79
	Asian	85	66	46.18	12.35	0.92	3.58
	Hispanic	287	66	38.84	12.39	0.91	3.79
	Black or African American	125	66	40.35	13.28	0.92	3.68
	White, Non-Hispanic	8458	66	44.11	11.95	0.91	3.65
	Native Hawaiian/Other Pacific Islander	31	66	45.32	13.83	0.93	3.56
	Female	4973	66	42.52	12.14	0.91	3.69
	Male	5289	66	42.98	12.77	0.92	3.69
	Limited English Proficient	347	66	28.55	10.90	0.88	3.71
	Migrant	21	66	37.90	12.12	0.90	3.81
	Plan 504	55	66	39.51	12.32	0.90	3.84
	All Students	10262	66	42.76	12.47	0.91	3.69
5	Special Education	1149	66	29.66	12.95	0.92	3.75
	Title 1	4300	66	37.60	12.75	0.91	3.72
	Low Income	4646	66	37.18	12.36	0.91	3.74

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
5	American Indian	1299	66	32.12	12.48	0.91	3.75
	Asian	91	66	42.71	13.34	0.93	3.57
	Hispanic	304	66	38.08	11.84	0.90	3.70
	Black or African American	141	66	38.20	13.50	0.92	3.73
	White, Non-Hispanic	8653	66	42.71	11.84	0.91	3.62
	Native Hawaiian/Other Pacific Islander	26	66	46.62	11.49	0.91	3.42
	Female	5045	66	41.10	12.06	0.91	3.66
	Male	5469	66	41.32	12.83	0.92	3.63
	Limited English Proficient	317	66	25.30	10.42	0.87	3.71
	Migrant	22	66	38.86	9.23	0.83	3.81
	Plan 504	45	66	37.58	12.79	0.91	3.80
	All Students	10514	66	41.22	12.47	0.91	3.65
6	Special Education	1136	66	25.50	10.70	0.89	3.58
	Title 1	3663	66	33.75	12.40	0.91	3.67
	Low Income	4467	66	33.52	12.10	0.91	3.69
	American Indian	1197	66	29.62	11.45	0.90	3.64
	Asian	94	66	42.24	13.28	0.93	3.56
	Hispanic	284	66	33.58	11.40	0.90	3.68
	Black or African American	118	66	32.45	11.02	0.88	3.74
	White, Non-Hispanic	8726	66	39.00	12.42	0.91	3.66
	Native Hawaiian/Other Pacific Islander	22	66	39.45	10.68	0.88	3.77
	Female	5062	66	37.73	12.43	0.91	3.68
	Male	5379	66	37.74	12.89	0.92	3.66
	Limited English Proficient	281	66	23.47	8.75	0.84	3.55
	Migrant	18	66	36.28	9.76	0.85	3.78
	Plan 504	67	66	36.09	12.49	0.91	3.70
	All Students	10442	66	37.73	12.67	0.92	3.67
7	Special Education	1106	66	23.03	10.35	0.88	3.58
	Title 1	3561	66	33.19	13.18	0.92	3.72
	Low Income	4317	66	32.61	12.77	0.91	3.73
	American Indian	1204	66	28.53	12.36	0.91	3.69
	Asian	94	66	42.36	13.42	0.93	3.59
	Hispanic	302	66	33.80	12.58	0.91	3.70
	Black or African American	124	66	32.15	12.31	0.91	3.71
	White, Non-Hispanic	8698	66	39.06	13.03	0.92	3.68

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
7	Native Hawaiian/Other Pacific Islander	28	66	36.96	14.51	0.94	3.67
	Female	5077	66	37.79	13.01	0.92	3.70
	Male	5373	66	37.49	13.76	0.93	3.69
	Limited English Proficient	291	66	21.34	9.49	0.86	3.53
	Migrant	12	66	31.58	7.53	0.77	3.63
	Plan 504	97	66	33.98	12.47	0.91	3.78
	All Students	10457	66	37.62	13.41	0.92	3.70
8	Special Education	1065	61	22.45	8.68	0.84	3.52
	Title 1	3528	61	30.02	10.61	0.89	3.57
	Low Income	4131	61	29.99	10.29	0.88	3.59
	American Indian	1149	61	26.60	10.08	0.88	3.56
	Asian	121	61	38.88	11.23	0.91	3.44
	Hispanic	310	61	30.61	10.17	0.88	3.59
	Black or African American	98	61	28.93	8.87	0.84	3.59
	White, Non-Hispanic	8908	61	35.19	10.70	0.89	3.52
	Native Hawaiian/Other Pacific Islander	36	61	32.92	9.94	0.88	3.51
	Female	5143	61	34.00	10.74	0.89	3.52
	Male	5479	61	34.20	11.19	0.90	3.53
	Limited English Proficient	231	61	20.38	7.88	0.81	3.44
	Migrant	17	61	34.35	11.71	0.91	3.49
	Plan 504	92	61	30.52	9.74	0.86	3.58
	All Students	10629	61	34.09	10.98	0.90	3.53
10	Special Education	964	66	19.69	7.96	0.81	3.47
	Title 1	2686	66	28.03	11.43	0.90	3.67
	Low Income	3412	66	27.56	10.93	0.89	3.66
	American Indian	1017	66	24.13	9.82	0.86	3.61
	Asian	105	66	36.07	12.15	0.91	3.65
	Hispanic	256	66	28.10	11.23	0.89	3.68
	Black or African American	111	66	26.90	11.00	0.89	3.62
	White, Non-Hispanic	8858	66	33.16	11.81	0.90	3.69
	Native Hawaiian/Other Pacific Islander	38	66	31.82	10.61	0.87	3.75
	Female	5053	66	31.63	11.58	0.90	3.70
	Male	5332	66	32.56	12.27	0.91	3.67
	Limited English Proficient	199	66	18.95	5.98	0.67	3.46
	Migrant	18	66	27.78	10.40	0.88	3.67
	Plan 504	108	66	29.83	11.30	0.89	3.69
	All Students	10389	66	32.10	11.95	0.90	3.69

Table L-2. 2009–10 MontCAS: Subgroup Reliabilities – Reading

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
3	Special Education	1169	60	30.61	11.46	0.92	3.34
	Title 1	4424	60	36.58	10.61	0.91	3.23
	Low Income	4884	60	36.17	10.43	0.90	3.24
	American Indian	1319	60	32.62	10.23	0.89	3.34
	Asian	94	60	42.07	10.68	0.92	3.00
	Hispanic	315	60	36.90	10.24	0.90	3.21
	Black or African American	149	60	37.00	10.39	0.90	3.25
	White, Non-Hispanic	8592	60	40.54	9.87	0.90	3.07
	Native Hawaiian/Other Pacific Islander	40	60	37.93	10.83	0.91	3.19
	Female	5156	60	40.38	9.97	0.90	3.09
	Male	5353	60	38.44	10.52	0.91	3.13
	Limited English Proficient	290	60	27.90	9.85	0.88	3.42
	Migrant	25	60	35.92	11.19	0.92	3.17
	Plan 504	39	60	40.05	8.73	0.87	3.15
	All Students	10509	60	39.39	10.29	0.91	3.11
4	Special Education	1195	60	30.23	11.24	0.91	3.40
	Title 1	4226	60	36.69	11.21	0.91	3.28
	Low Income	4662	60	36.73	10.99	0.91	3.28
	American Indian	1270	60	32.61	11.23	0.91	3.39
	Asian	84	60	42.71	10.49	0.92	3.03
	Hispanic	287	60	36.24	11.02	0.91	3.30
	Black or African American	124	60	38.19	11.27	0.92	3.23
	White, Non-Hispanic	8429	60	41.25	10.01	0.91	3.08
	Native Hawaiian/Other Pacific Islander	31	60	41.26	12.15	0.94	3.01
	Female	4962	60	40.76	10.37	0.91	3.12
	Male	5263	60	39.30	10.83	0.92	3.13
	Limited English Proficient	346	60	25.86	9.76	0.87	3.51
	Migrant	21	60	35.90	11.64	0.91	3.40
	Plan 504	54	60	39.09	10.45	0.90	3.23
	All Students	10225	60	40.01	10.63	0.91	3.13
5	Special Education	1130	60	31.10	11.23	0.91	3.41
	Title 1	4333	60	38.58	10.73	0.91	3.22
	Low Income	4632	60	38.24	10.66	0.91	3.23
	American Indian	1292	60	33.87	10.89	0.90	3.36
	Asian	91	60	42.99	10.22	0.91	3.00

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
5	Hispanic	301	60	39.25	10.61	0.91	3.21
	Black or African American	140	60	40.14	10.55	0.91	3.13
	White, Non-Hispanic	8641	60	42.94	9.45	0.90	3.03
	Native Hawaiian/Other Pacific Islander	26	60	44.69	6.72	0.79	3.05
	Female	5037	60	42.74	9.77	0.90	3.04
	Male	5454	60	40.71	10.38	0.91	3.11
	Limited English Proficient	315	60	26.45	9.79	0.88	3.45
	Migrant	22	60	37.64	8.43	0.85	3.25
	Plan 504	45	60	38.24	11.37	0.92	3.22
	All Students	10491	60	41.68	10.15	0.91	3.09
6	Special Education	1117	60	29.26	10.77	0.90	3.46
	Title 1	3619	60	37.61	10.93	0.91	3.33
	Low Income	4455	60	37.61	10.75	0.90	3.33
	American Indian	1192	60	33.39	10.92	0.90	3.43
	Asian	94	60	43.15	10.12	0.90	3.14
	Hispanic	283	60	38.14	10.53	0.90	3.34
	Black or African American	119	60	38.83	9.21	0.87	3.30
	White, Non-Hispanic	8708	60	42.59	9.70	0.90	3.13
	Native Hawaiian/Other Pacific Islander	22	60	41.41	11.10	0.92	3.10
	Female	5057	60	42.50	9.89	0.90	3.12
	Male	5361	60	40.31	10.58	0.91	3.19
	Limited English Proficient	280	60	25.13	9.17	0.85	3.55
	Migrant	18	60	37.72	9.88	0.89	3.27
	Plan 504	67	60	41.16	8.76	0.86	3.28
	All Students	10419	60	41.37	10.32	0.91	3.17
7	Special Education	1109	60	27.75	10.84	0.90	3.44
	Title 1	3480	60	37.94	11.73	0.92	3.25
	Low Income	4318	60	37.49	11.46	0.92	3.27
	American Indian	1206	60	33.72	11.99	0.92	3.37
	Asian	93	60	43.89	10.65	0.92	2.99
	Hispanic	301	60	38.75	11.14	0.92	3.21
	Black or African American	122	60	39.46	10.97	0.91	3.24
	White, Non-Hispanic	8704	60	42.79	10.28	0.91	3.05
	Native Hawaiian/Other Pacific Islander	27	60	41.41	8.01	0.83	3.30
	Female	5080	60	43.13	10.43	0.92	3.04
	Male	5373	60	40.13	11.17	0.92	3.13

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
7	Limited English Proficient	290	60	25.05	9.75	0.87	3.51
	Migrant	12	60	33.25	11.79	0.93	3.23
	Plan 504	97	60	39.87	11.31	0.92	3.20
	All Students	10460	60	41.58	10.93	0.92	3.10
8	Special Education	1073	60	29.61	10.83	0.90	3.49
	Title 1	3438	60	39.08	11.36	0.92	3.31
	Low Income	4128	60	39.00	11.21	0.91	3.30
	American Indian	1148	60	35.23	11.89	0.92	3.40
	Asian	121	60	45.91	9.94	0.91	2.97
	Hispanic	311	60	39.95	10.74	0.91	3.27
	Black or African American	97	60	42.39	8.64	0.86	3.27
	White, Non-Hispanic	8917	60	44.14	9.92	0.90	3.09
	Native Hawaiian/Other Pacific Islander	36	60	44.92	8.56	0.87	3.12
	Female	5154	60	44.69	9.95	0.90	3.07
	Male	5476	60	41.53	10.86	0.92	3.16
	Limited English Proficient	230	60	25.67	10.25	0.88	3.51
	Migrant	17	60	41.71	11.03	0.92	3.09
	Plan 504	92	60	40.22	9.23	0.87	3.33
	All Students	10637	60	43.05	10.56	0.91	3.13
10	Special Education	990	60	29.43	10.39	0.89	3.49
	Title 1	2480	60	39.03	11.35	0.92	3.22
	Low Income	3430	60	38.99	11.05	0.91	3.24
	American Indian	1010	60	35.65	11.31	0.91	3.36
	Asian	107	60	43.62	8.53	0.87	3.05
	Hispanic	256	60	38.75	10.85	0.91	3.28
	Black or African American	111	60	40.05	10.31	0.90	3.18
	White, Non-Hispanic	8892	60	43.56	10.05	0.91	3.02
	Native Hawaiian/Other Pacific Islander	38	60	44.50	8.49	0.88	2.97
	Female	5064	60	44.55	9.47	0.90	2.97
	Male	5350	60	40.84	11.04	0.92	3.12
	Limited English Proficient	195	60	28.21	9.26	0.85	3.55
	Migrant	18	60	38.78	11.97	0.92	3.30
	Plan 504	110	60	42.62	9.20	0.88	3.14
	All Students	10422	60	42.63	10.48	0.91	3.07

Table L-3. 2009–10 MontCAS: Subgroup Reliabilities – Science

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
4	Special Education	1236	61	32.20	10.02	0.87	3.64
	Title 1	8	61				
	Low Income	4689	61	35.96	9.55	0.86	3.58
	American Indian	1276	61	31.87	9.48	0.85	3.65
	Asian	85	61	40.64	9.56	0.87	3.38
	Hispanic	288	61	35.88	9.15	0.85	3.58
	Black or African American	125	61	36.18	9.99	0.88	3.53
	White, Non-Hispanic	8460	61	40.13	8.89	0.85	3.43
	Native Hawaiian/Other Pacific Islander	31	61	40.06	10.24	0.88	3.58
	Female	4975	61	38.79	9.29	0.86	3.47
	Male	5290	61	39.08	9.54	0.87	3.47
	Limited English Proficient	350	61	26.70	8.00	0.79	3.68
	Migrant	21	61	33.90	8.90	0.84	3.59
	Plan 504	56	61	38.45	9.40	0.86	3.52
	All Students	10265	61	38.94	9.42	0.86	3.48
8	Special Education	1114	61	26.76	10.20	0.86	3.83
	Title 1	18	61	32.22	9.96	0.87	3.66
	Low Income	4154	61	32.80	10.66	0.87	3.79
	American Indian	1149	61	28.46	9.98	0.85	3.81
	Asian	121	61	39.56	10.04	0.87	3.63
	Hispanic	314	61	33.23	10.19	0.86	3.79
	Black or African American	98	61	33.42	9.80	0.85	3.78
	White, Non-Hispanic	8946	61	38.20	10.34	0.87	3.67
	Native Hawaiian/Other Pacific Islander	36	61	34.39	9.93	0.86	3.74
	Female	5158	61	36.09	10.35	0.87	3.71
	Male	5506	61	37.78	11.04	0.89	3.68
	Limited English Proficient	231	61	21.85	7.54	0.76	3.71
	Migrant	19	61	34.53	12.17	0.91	3.66
	Plan 504	92	61	35.77	10.72	0.88	3.75
	All Students	10671	61	36.96	10.75	0.88	3.70
10	Special Education	1023	61	22.90	8.90	0.84	3.52
	Title 1	102	61	30.39	10.46	0.88	3.60
	Low Income	3442	61	30.12	10.32	0.88	3.57
	American Indian	1011	61	26.31	9.51	0.86	3.58
	Asian	107	61	35.52	9.83	0.87	3.49

continued

Grade	Group	Number of students	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
10	Hispanic	259	61	30.08	10.56	0.89	3.58
	Black or African American	111	61	30.05	10.17	0.88	3.57
	White, Non-Hispanic	8916	61	35.01	10.41	0.89	3.51
	Native Hawaiian/Other Pacific Islander	38	61	34.76	8.38	0.82	3.54
	Female	5070	61	33.48	10.07	0.88	3.54
	Male	5372	61	34.48	11.15	0.90	3.49
	Limited English Proficient	197	61	20.55	6.66	0.73	3.47
	Migrant	19	61	30.11	11.38	0.90	3.52
	Plan 504	110	61	33.51	9.93	0.87	3.59
	All Students	10450	61	33.99	10.65	0.89	3.52

Table L-4. 2009–10 MontCAS: Reliabilities by Reporting Category -- Mathematics

Grade	Item reporting category	Number of items	Raw score			Alpha	SEM
			Maximum	Mean	Standard deviation		
3	2	19	22	15.51	3.81	0.77	1.83
	3	8	8	5.31	1.84	0.61	1.15
	4	7	10	5.45	2.13	0.54	1.44
	5	10	10	7.18	1.86	0.55	1.24
	6	8	8	5.92	1.89	0.66	1.10
	7	8	8	5.76	1.73	0.61	1.09
4	2	19	22	12.57	5.02	0.82	2.10
	3	8	8	5.23	2.00	0.65	1.18
	4	10	10	7.09	1.81	0.48	1.30
	5	10	10	7.37	1.95	0.62	1.20
	6	8	8	5.93	1.82	0.62	1.12
	7	5	8	4.57	2.43	0.47	1.78
5	2	18	21	11.65	4.90	0.82	2.11
	3	8	8	5.61	1.87	0.65	1.11
	4	11	11	7.36	2.11	0.54	1.42
	5	8	8	4.70	1.83	0.56	1.21
	6	10	10	6.91	2.32	0.68	1.31
	7	5	8	4.98	1.95	0.39	1.52
6	2	18	21	10.82	4.88	0.81	2.13
	3	8	8	4.97	1.90	0.60	1.21
	4	11	11	7.47	2.13	0.57	1.39
	5	8	8	4.76	1.86	0.56	1.24
	6	7	10	4.76	2.50	0.61	1.55
	7	8	8	4.95	1.91	0.60	1.20
7	2	18	18	9.61	4.22	0.81	1.82
	3	8	8	4.91	2.10	0.68	1.19
	4	9	12	6.37	2.81	0.63	1.70
	5	8	8	4.49	1.78	0.51	1.25
	6	9	12	7.10	2.92	0.64	1.74
	7	8	8	5.13	1.97	0.64	1.18
8	2	14	17	8.96	3.47	0.66	2.03
	3	8	8	4.67	2.21	0.71	1.18
	4	8	8	4.49	1.67	0.46	1.23
	5	8	8	4.09	1.78	0.52	1.23
	6	12	12	6.91	2.56	0.65	1.52
	7	8	8	4.97	1.88	0.58	1.23
10	2	13	13	6.54	3.10	0.74	1.57
	3	11	11	5.96	2.46	0.65	1.45
	4	10	13	5.25	2.56	0.60	1.62
	5	8	8	3.63	1.85	0.53	1.26
	6	10	13	6.54	2.74	0.55	1.83
	7	8	8	4.18	1.82	0.56	1.21

Table L-5. 2009–10 MontCAS: Reliabilities by Reporting Category – Reading

Grade	*Item reporting category*	*Number of items*	*Raw score*			*Alpha*	*SEM*
			Maximum	*Mean*	*Standard deviation*		
3	1	20	20	14.19	3.88	0.79	1.79
	2	18	18	13.34	3.48	0.78	1.63
	4	9	12	6.74	2.37	0.62	1.45
	5	7	10	5.12	1.94	0.54	1.31
4	1	15	18	11.39	3.45	0.76	1.70
	2	19	19	14.05	3.63	0.78	1.71
	4	9	12	6.83	2.40	0.63	1.45
	5	11	11	7.75	2.50	0.71	1.36
5	1	18	21	14.73	3.76	0.76	1.84
	2	18	18	12.97	3.63	0.78	1.70
	4	9	9	5.92	1.99	0.61	1.24
	5	9	12	8.06	2.07	0.60	1.30
6	1	20	20	15.17	3.67	0.78	1.71
	2	17	20	12.59	3.81	0.75	1.90
	4	9	9	6.77	1.83	0.61	1.15
	5	8	11	6.83	2.34	0.60	1.48
7	1	15	18	12.43	3.27	0.73	1.69
	2	17	20	13.56	4.06	0.81	1.79
	4	9	9	6.51	1.98	0.65	1.17
	5	13	13	9.09	2.82	0.72	1.48
8	1	18	21	14.49	3.85	0.74	1.95
	2	18	18	13.28	3.59	0.78	1.68
	4	10	13	9.14	2.42	0.63	1.47
	5	8	8	6.15	1.91	0.70	1.04
10	1	13	16	10.85	2.99	0.71	1.60
	2	17	17	12.44	3.44	0.78	1.63
	4	12	15	9.99	2.87	0.69	1.59
	5	12	12	9.35	2.41	0.71	1.30

Table L-6. 2009–10 MontCAS: Reliabilities by Reporting Category – Science

Grade	*Item reporting category*	*Number of items*	*Raw score*			*Alpha*	*SEM*
			Maximum	*Mean*	*Standard deviation*		
4	1	14	14	9.18	2.64	0.65	1.57
	2	14	14	9.75	2.52	0.59	1.60
	3	11	14	8.59	2.75	0.52	1.90
	4	11	14	7.97	2.44	0.57	1.61
	5	2	2	1.64	0.61	0.39	0.47
	6	3	3	1.82	0.93	0.29	0.79
8	1	11	14	9.13	2.98	0.68	1.69
	2	14	14	8.17	2.70	0.63	1.64
	3	11	14	8.17	3.28	0.59	2.11
	4	14	14	8.13	2.75	0.63	1.67
	5	3	3	1.87	0.86	0.27	0.74
	6	2	2	1.49	0.67	0.34	0.55
10	1	11	14	8.01	3.15	0.68	1.78
	2	11	14	6.43	2.73	0.66	1.58
	3	14	14	8.00	2.83	0.65	1.68
	4	14	14	8.35	2.88	0.67	1.66
	5	1	1	0.53	0.50	0.67	0.29
	6	4	4	2.66	1.09	0.35	0.88

Appendix M—Interrater Agreement

Table M-1. 2009–10 MontCAS: Item Level Interrater Consistency Statistics by Subject and Grade

Subject	*Grade*	*Item*	*Number of score categories*	*Number of responses scored twice*	*Percent exact*	*Percent adjacent*	*Correlation*	*Percent of third scores*
Mathematics	3	42994	2	214	100.00	0.00	1.00	0.00
		59293	2	213	99.53	0.47	0.98	0.00
		43261	5	214	90.19	8.88	0.92	0.93
		59294	2	212	98.58	1.42	0.96	0.00
		76899	5	208	93.75	6.25	0.97	0.00
	4	34900	2	217	98.16	1.84	0.96	0.00
		61780	2	211	100.00	0.00	1.00	0.00
		77063	5	210	89.05	9.52	0.95	1.43
		76763	2	210	99.05	0.95	0.98	0.00
		62486	5	222	88.29	9.46	0.97	2.25
	5	77298	2	213	98.59	1.41	0.97	0.00
		62025	2	210	98.10	1.90	0.96	0.00
		61052	5	210	81.43	15.71	0.91	2.86
		77295	2	213	99.53	0.47	0.98	0.00
		34660	5	208	90.38	8.65	0.97	0.96
	6	63005	2	214	98.13	1.87	0.96	0.00
		44088	2	213	96.24	3.76	0.85	0.00
		77642	2	214	98.13	1.87	0.96	0.00
		77649	5	217	79.72	16.13	0.90	4.15
		44048	5	206	79.61	19.42	0.92	0.97
	7	86348	2	213	98.12	1.88	0.96	0.00
		43799	2	213	97.65	2.35	0.94	0.00
		86349	2	205	99.02	0.98	0.98	0.00
		15863	5	197	80.20	17.26	0.92	2.54
		43922	5	213	85.92	13.15	0.96	0.47
	8	44199	2	212	99.06	0.94	0.98	0.00
		87834	2	212	99.53	0.47	0.99	0.00
		87841	2	216	97.69	2.31	0.95	0.00
		44267	5	207	75.36	21.74	0.91	2.90
	10	43899	2	214	100.00	0.00	1.00	0.00
		34864	2	192	88.54	11.46	0.76	0.00
		77635	2	211	98.58	1.42	0.97	0.00
		16052	5	197	63.96	29.44	0.86	5.08
		15964	5	204	78.43	20.10	0.86	0.98
Reading	3	92761	5	217	61.29	36.41	0.65	1.84
		92783	5	211	71.56	26.07	0.78	2.37
	4	67382	5	207	63.29	36.23	0.79	0.48
		94139	5	215	66.98	31.63	0.76	1.40
	5	93668	5	211	64.45	33.18	0.79	2.37
		93462	5	217	70.97	29.03	0.72	0.00
	6	95397	5	213	69.95	27.70	0.83	1.88
		95171	5	217	75.12	23.96	0.83	0.92
	7	92611	5	208	60.10	37.98	0.74	1.92
		92562	5	209	63.16	36.36	0.78	0.48
	8	95869	5	209	53.11	43.54	0.72	1.91
		68511	5	217	59.45	37.33	0.77	2.76
	10	94887	5	206	64.56	34.95	0.77	0.49
		94955	5	211	65.40	33.18	0.77	1.42
Science	4	42790	5	194	85.05	14.43	0.95	0.52
		75427	5	207	85.02	13.53	0.87	1.45

continued

Subject	*Grade*	*Item*	*Number of score categories*	*Number of responses scored twice*	*Percent exact*	*Percent adjacent*	*Correlation*	*Percent of third scores*
Science	8	55106	5	217	68.20	29.49	0.84	2.30
		39764	5	223	75.78	16.59	0.91	7.62
	10	75461	5	227	66.96	28.19	0.85	5.29
		75652	5	201	78.61	20.40	0.79	1.00

Appendix N—Decision Accuracy and Consistency Results

Table N-1. 2009–10 MontCAS: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—Overall and Conditional on Performance Level

Subject	*Grade*	*Overall*	*Kappa*	*Conditional on level*			
				Novice	*Nearing Proficiency*	*Proficient*	*Advanced*
Mathematics	3	0.78 (0.70)	0.58	0.82 (0.74)	0.62 (0.51)	0.77 (0.70)	0.89 (0.80)
	4	0.79 (0.71)	0.59	0.81 (0.73)	0.62 (0.51)	0.77 (0.69)	0.90 (0.82)
	5	0.79 (0.71)	0.59	0.81 (0.73)	0.63 (0.52)	0.76 (0.69)	0.90 (0.82)
	6	0.80 (0.72)	0.60	0.82 (0.74)	0.62 (0.51)	0.78 (0.71)	0.90 (0.83)
	7	0.79 (0.71)	0.60	0.80 (0.72)	0.60 (0.49)	0.74 (0.66)	0.92 (0.86)
	8	0.78 (0.70)	0.57	0.79 (0.68)	0.67 (0.57)	0.77 (0.71)	0.88 (0.79)
	10	0.77 (0.68)	0.55	0.65 (0.52)	0.72 (0.64)	0.77 (0.69)	0.89 (0.80)
Reading	3	0.85 (0.79)	0.65	0.75 (0.61)	0.75 (0.66)	0.81 (0.75)	0.91 (0.86)
	4	0.84 (0.78)	0.64	0.77 (0.66)	0.72 (0.62)	0.80 (0.74)	0.91 (0.86)
	5	0.85 (0.79)	0.64	0.75 (0.59)	0.71 (0.60)	0.78 (0.72)	0.92 (0.88)
	6	0.85 (0.79)	0.64	0.80 (0.69)	0.65 (0.53)	0.79 (0.72)	0.93 (0.88)
	7	0.85 (0.79)	0.66	0.81 (0.71)	0.71 (0.61)	0.80 (0.74)	0.92 (0.88)
	8	0.84 (0.79)	0.64	0.80 (0.69)	0.66 (0.54)	0.76 (0.69)	0.93 (0.89)
	10	0.82 (0.75)	0.61	0.83 (0.74)	0.58 (0.46)	0.77 (0.70)	0.91 (0.85)
Science	4	0.78 (0.69)	0.53	0.75 (0.59)	0.75 (0.67)	0.79 (0.73)	0.83 (0.67)
	8	0.77 (0.69)	0.54	0.76 (0.64)	0.73 (0.64)	0.78 (0.71)	0.84 (0.70)
	10	0.74 (0.65)	0.53	0.82 (0.75)	0.72 (0.64)	0.64 (0.54)	0.86 (0.73)

Table N-2. 2009–10 MontCAS: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—Conditional on Cutpoint

Subject	Grade	Novice/Nearing Proficiency			Nearing Proficiency/Proficient			Proficient/Advanced		
		Accuracy (consistency)	*False positive*	*False negative*	*Accuracy (consistency)*	*False positive*	*False negative*	*Accuracy (consistency)*	*False positive*	*False negative*
Mathematics	3	0.95 (0.93)	0.02	0.03	0.92 (0.88)	0.04	0.04	0.91 (0.88)	0.06	0.03
	4	0.95 (0.93)	0.02	0.03	0.92 (0.89)	0.04	0.03	0.91 (0.88)	0.06	0.03
	5	0.95 (0.94)	0.02	0.02	0.92 (0.89)	0.04	0.04	0.91 (0.88)	0.06	0.03
	6	0.95 (0.93)	0.02	0.02	0.92 (0.89)	0.04	0.04	0.92 (0.89)	0.05	0.03
	7	0.94 (0.92)	0.03	0.03	0.92 (0.89)	0.04	0.03	0.92 (0.89)	0.05	0.03
	8	0.96 (0.94)	0.02	0.02	0.91 (0.88)	0.05	0.04	0.91 (0.88)	0.06	0.03
	10	0.93 (0.90)	0.03	0.04	0.91 (0.87)	0.05	0.04	0.93 (0.91)	0.04	0.02
Reading	3	0.99 (0.98)	0.01	0.01	0.95 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04
	4	0.98 (0.97)	0.01	0.01	0.95 (0.93)	0.02	0.02	0.91 (0.87)	0.05	0.04
	5	0.99 (0.98)	0	0.01	0.95 (0.93)	0.02	0.03	0.91 (0.87)	0.06	0.04
	6	0.98 (0.97)	0.01	0.01	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04
	7	0.98 (0.97)	0.01	0.01	0.96 (0.94)	0.02	0.02	0.91 (0.88)	0.05	0.04
	8	0.98 (0.97)	0.01	0.01	0.95 (0.93)	0.02	0.02	0.91 (0.88)	0.05	0.04
	10	0.97 (0.96)	0.01	0.02	0.95 (0.92)	0.03	0.03	0.90 (0.86)	0.06	0.04
Science	4	0.97 (0.96)	0.01	0.02	0.90 (0.85)	0.06	0.05	0.91 (0.88)	0.06	0.02
	8	0.96 (0.94)	0.02	0.02	0.90 (0.87)	0.05	0.04	0.91 (0.88)	0.06	0.03
	10	0.92 (0.89)	0.04	0.04	0.90 (0.85)	0.06	0.04	0.93 (0.90)	0.05	0.02

Appendix O—REPORT SHELLS

MontCAS
Criterion-Referenced Test (CRT)
Student Report
2010

Letter from Superintendent

Dear Parents/Guardians:

The Montana Comprehensive Assessment System (MontCAS) Criterion-Referenced Test (CRT) is the state's measure of student performance on the state content standards which establish goals for what all students should know and be able to do.

The CRT assesses Reading and Math at grades 3-8 and 10. Students in grades 4, 8, and 10 are also assessed in Science. The assessment contains multiple-choice questions, math short answer questions, and constructed response items. The constructed response items give students the opportunity to explain answers and solve problems using multiple strategies. This report shows how your student performed on the March 2010 CRT.

The results of this standards-based assessment are reported in four performance levels: Advanced, Proficient, Nearing Proficiency, and Novice. While some students may not yet meet the standards, keep in mind that the standards are rigorous and challenging. Our long term goal is for all students to achieve these high standards so that Montana youth will be among the best educated in the world. The staff at your school will be able to provide further information about your student's performance on the CRT.

The CRT is only one measure of student performance and should be viewed in the context of the student's local programs and other measures. The CRT is required by the No Child Left Behind Act and is part of an ongoing statewide educational improvement process. I encourage you to contact your student's school to begin a conversation that will support your student's success.

Sincerely,

Denise Juneau

Denise Juneau
Montana Superintendent of Public Instruction
Montana Office of Public Instruction
PO Box 202501
Helena, Montana 59620-2501
http://www.opi.mt.gov

What can you do to help your student?

It is important to support your student in his or her studies now and throughout his or her future education.

Here are some tips for supporting your student in the completion of his or her schoolwork:

- Have regular discussions with your student's teacher(s) to see what you can do at home to support your student's work in school, such as making sure homework is done.
- Discuss with your student the subjects in which he or she needs improvement. Talk about whether there has been a noticeable improvement. If not, find out why.
- Ask your student to explain what he or she is studying. These conversations help you to follow your student's progress and help your student to remember what he or she has learned.
- Make sure your student gets enough rest, eats properly, and arrives at school on time every day. Send your student to school prepared to learn.

What is the MontCAS Criterion-Referenced Test (CRT)?

The Montana Comprehensive Assessment System (MontCAS) was developed in accordance with the following federal laws: Title 1 of the Elementary and Secondary Education Act (ESEA) of 1994, P. L. 103-382, and the No Child Left Behind Act (NCLB) of 2001.

The CRT test questions are based on, and aligned to, Montana's content standards, benchmarks, and grade-level expectations in Mathematics, Reading, and Science. Montana educators worked with the Montana Office of Public Instruction and Measured Progress to develop test questions that assess how well students have met Montana grade-level expectations for each content

MontCAS CRT scores are intended to be useful indicators of the extent to which students have mastered the materials outlined in the Montana Mathematics, Reading, and Science content standards, benchmarks, and grade-level expectations.

Who must take the CRT?

All classroom students in grades 3-8 and 10 enrolled for 180 hours or more in an accredited public or private Montana school are required to participate.

What subjects were tested in spring 2010?

Mathematics	Grades 3-8 and 10
Reading	Grades 3-8 and 10
Science	Grades 4, 8, and 10

What types of test questions are on the CRT?

- Multiple-choice questions: Students choose the correct answer from four options and receive one point for each correct answer and zero points for an incorrect answer.
- Constructed-response questions: Depending on the subject tested, students are asked to explain and/or make a chart, table, diagram, illustration, or graph to support their answer. Each answer receives zero to four points.
- Short-answer questions (Mathematics tests only): Students give a brief response, which is usually a number or short statement. Students receive one point for a correct answer and zero points for an incorrect answer.

How are the CRT results used?

MontCAS CRT test results are used for the following purposes:

- to assist educators in planning improvements to curriculum and instruction
- to determine whether schools are helping their students meet the state content standards

Where can you find more information?

Montana test results for all schools and districts in the state:
http://opi.mt.gov/Curriculum/MontCAS/index.html

Montana requirements for the participation of students with disabilities on the CRT: http://opi.mt.gov/Curriculum/MontCAS/#gpm1_7

OPI contact: Judy Snow, State Assessment Director, 406-444-3656, jsnow@mt.gov

Name: SCHREFFLER, JACOB
SASID: D04100027

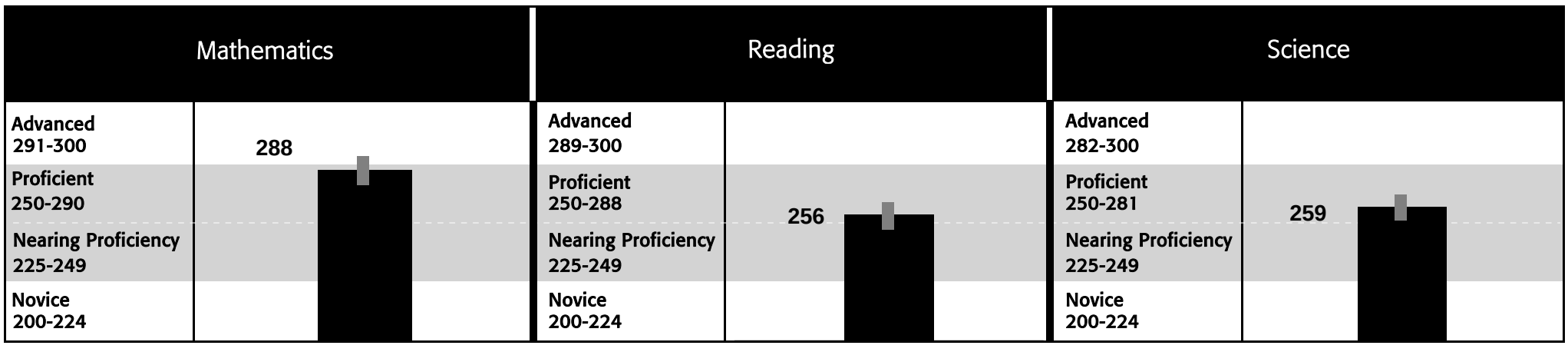
Date of Birth: 06/23/2000
Grade: 04

School: Demonstration School 1
System: Demonstration District A

Your student's performance level and score in each content area

Display of scores and probable range of scores

In the figure below, the top of the black bar indicates your student's score on each test. The smaller gray bar shows the range of likely scores your student could have received if he or she had taken the test multiple times.



Your student's Mathematics Scaled Score is **288** which is at the **Proficient Level.** Your student's possible range of scores is from 278 to 298.

This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

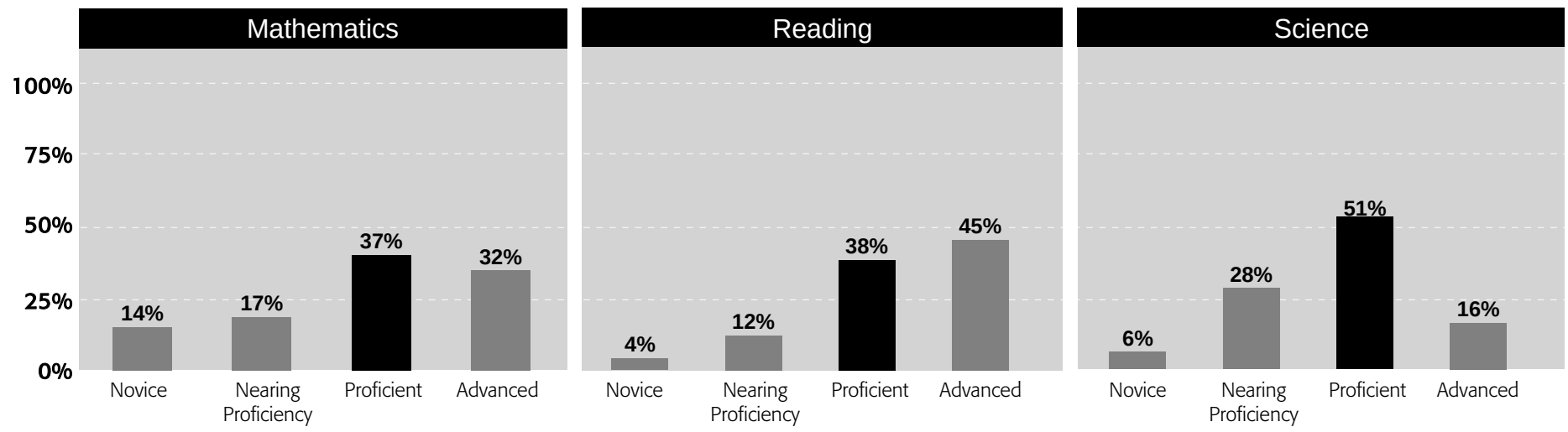
Your student's Reading Scaled Score is **256** which is at the **Proficient Level.** Your student's possible range of scores is from 247 to 265.

This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Your student's Science Scaled Score is **259** which is at the **Proficient Level.** Your student's possible range of scores is from 252 to 266.

This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Your student's performance level compared to the State



Black bar indicates the percent of students in your student's performance level for each subject

Scores on Montana Content Standards

CRT results are reported for Montana Content Standards in Mathematics, Reading, and Science to provide standard-specific information about the student's achievement. The results can be used to show the student's relative performance on the standards within a content area.

Mathematics	Total Possible Points	Points Earned by Your Student	Range of Points Earned by Students Who Have Achieved Proficiency in the State
1. Problem Solving	This standard is assessed within the frameworks of standard 2-7.		
2. Numbers and Operations	22	13	5-22
3. Algebra	8	6	0-8
4. Geometry	10	6	2-10
5. Measurement	10	10	2-10
6. Data Analysis, Statistics, and Probability	8	7	2-8
7. Patterns, Relations, and Functions	8	8	0-8

Reading	Total Possible Points	Points Earned by Your Student	Range of Points Earned by Students Who Have Achieved Proficiency in the State
1. Students construct meaning as they comprehend, interpret, and respond to what they read.	18	8	2-16
2. Students apply a range of skills and strategies to read.	19	10	6-19
3. Students set goals, monitor, and evaluate their reading progress.	This standard is not measurable in a statewide assessment.		
4. Students select, read, and respond to print and nonprint material for a variety of purposes.	12	7	1-10
5. Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences.	11	6	2-11

Science	Total Possible Points	Points Earned by Your Student	Range of Points Earned by Students Who Have Achieved Proficiency in the State
1. Scientific Investigations	14	10	4-14
2. Physical Science	14	10	3-14
3. Life Science	14	8	3-14
4. Earth/Space Science	14	7	2-14
5. Impact on Society	Subscores are not reported for this standard.		
6. Historical Development	Subscores are not reported for this standard.		

MontCAS CRT

School: Demonstration School 1
System: Demonstration District A
Grade: 03
Spring 2010

Mathematics

School Summary Report

I. Distribution of Scores

Perf. Level	Scores	School			System			State		
		N	% of Students	% of Students in Cat.	N	% of Students	% of Students in Cat.	N	% of Students	% of Students in Cat.
Advanced	299-300	5	25	35	8	22	35	1926	18	29
	297-298	0	0		2	5		370	4	
	294-296	0	0		1	3		373	4	
	292-293	2	10		2	5		380	4	
	290-291	0	0		0	0		0	0	
Proficient	282-289	3	15	35	3	8	27	1199	11	40
	274-281	2	10		3	8		765	7	
	266-273	0	0		1	3		1074	10	
	258-265	2	10		3	8		622	6	
	250-257	0	0		0	0		610	6	
Nearing Proficiency	245-249	2	10	20	3	8	16	526	5	16
	240-244	0	0		0	0		239	2	
	235-239	2	10		2	5		425	4	
	230-234	0	0		0	0		370	4	
	225-229	0	0		1	3		143	1	
Novice	220-224	1	5	10	2	5	22	307	3	14
	215-219	0	0		1	3		140	1	
	210-214	0	0		1	3		255	2	
	205-209	0	0		0	0		103	1	
	200-204	1	5		4	11		718	7	

II. Subtest Results

Mathematics		Possible Points	Average Points Earned		
			School	System	State
Total Points		66	47	44	45
Standards	1. Problem Solving	This standard is assessed within the frameworks of standards 2–7.			
	2. Numbers and Operations	22	16	15	16
	3. Algebra	8	6	5	5
	4. Geometry	10	6	6	5
	5. Measurement	10	7	7	7
	6. Data Analysis, Statistics, and Probability	8	6	6	6
	7. Patterns, Relations, and Functions	8	6	5	6

CRT Performance Level Descriptors

Advanced (290-300)
This level denotes superior performance.

Proficient (250-289)
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Nearing Proficiency (225-249)
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

Novice (200-224)
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

MontCAS CRT

Confidential

Mathematics

School Summary Report

School: Demonstration School 1
System: Demonstration District A
Grade: 03
Spring 2010

III. Results for Subgroups of Students

Reporting Category	School					System					State				
	Number	% in N	% in NP	% in P	% in A	Number	% in N	% in NP	% in P	% in A	Number	% in N	% in NP	% in P	% in A
All Students	20	10	20	35	35	37	22	16	27	35	10545	14	16	40	29
Gender															
Male	11	18	27	27	27	20	25	25	25	25	5378	15	16	40	30
Female	8	*	*	*	*	16	19	6	25	50	5167	14	17	41	28
Ethnicity															
American Indian or Alaska Native	3	*	*	*	*	7	*	*	*	*	1317	35	24	30	11
Asian	0	*	*	*	*	0	*	*	*	*	94	10	10	35	46
Hispanic	2	*	*	*	*	3	*	*	*	*	318	19	18	41	22
Black or African American	1	*	*	*	*	3	*	*	*	*	150	21	22	41	16
Native Hawaiian or Other Pacific Islander	0	*	*	*	*	1	*	*	*	*	40	15	23	33	30
White	12	8	17	42	33	21	5	19	33	43	8626	11	15	42	32
Special Education	3	*	*	*	*	6	*	*	*	*	1198	41	20	27	12
Students with a 504 Plan	0	*	*	*	*	1	*	*	*	*	39	8	23	38	31
Title I (optional)	6	*	*	*	*	15	20	27	13	40	4096	20	19	39	22
Tested with Standard Accommodation	4	*	*	*	*	6	*	*	*	*	1847	33	23	34	10
Tested with Non-Standard Accommodation	0	*	*	*	*	1	*	*	*	*	34	38	38	21	3
Alternate Assessment	If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report														
Migrant	1	*	*	*	*	1	*	*	*	*	25	28	24	36	12
Gifted/Talented	1	*	*	*	*	2	*	*	*	*	419	0	3	18	79
LEP/ELL	1	*	*	*	*	1	*	*	*	*	291	50	24	24	2
Former LEP Student	1	*	*	*	*	1	*	*	*	*	211	28	23	37	11
LEP Student Enrolled for First Time in a U.S. School	0	Performance levels are not reported for 1st year LEP students													
Free/Reduced Lunch	6	*	*	*	*	18	39	22	17	22	4907	21	20	40	19

*Less than ten (10) students were assessed

MontCAS CRT

System: Demonstration District A
Grade: 04
Spring 2010

Mathematics

System Summary Report

I. Distribution of Scores

Perf. Level	Scores	System			State		
		Number	% of Students	% of Students in Cat.	Number	% of Students	% of Students in Cat.
Advanced	299-300	6	18	24	2365	23	32
	297-298	0	0		295	3	
	295-296	1	3		318	3	
	293-294	0	0		0	0	
	291-292	1	3		303	3	
Proficient	283-290	7	21	47	902	9	37
	275-282	1	3		596	6	
	266-274	2	6		1050	10	
	258-265	2	6		516	5	
	250-257	4	12		764	7	
Nearing Proficiency	245-249	2	6	18	462	5	17
	240-244	2	6		388	4	
	235-239	0	0		183	2	
	230-234	2	6		374	4	
	225-229	0	0		338	3	
Novice	220-224	0	0	12	150	1	14
	215-219	3	9		293	3	
	210-214	0	0		228	2	
	205-209	0	0		97	1	
	200-204	1	3		641	6	

II. Subtest Results

Mathematics		Possible Points	Average Points Earned	
			System	State
Total Points		66	43	43
Standards	1. Problem Solving	This standard is assessed within the frameworks of standards 2–7.		
	2. Numbers and Operations	22	13	13
	3. Algebra	8	5	5
	4. Geometry	10	7	7
	5. Measurement	10	7	7
	6. Data Analysis, Statistics, and Probability	8	6	6
	7. Patterns, Relations, and Functions	8	4	5

CRT Performance Level Descriptors

Advanced (291-300)
This level denotes superior performance.

Proficient (250-290)
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Nearing Proficiency (225-249)
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

Novice (200-224)
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

MontCAS CRT

Confidential

Mathematics

System Summary Report

System: Demonstration District A
Grade: 04
Spring 2010

III. Results for Subgroups of Students

Reporting Category	System					State				
	Number	% in N	% in NP	% in P	% in A	Number	% in N	% in NP	% in P	% in A
All Students	34	12	18	47	24	10263	14	17	37	32
Gender										
Male	18	11	17	56	17	5290	14	17	35	34
Female	16	13	19	38	31	4973	13	17	39	30
Ethnicity										
American Indian or Alaska Native	3	*	*	*	*	1276	33	23	32	12
Asian	1	*	*	*	*	85	8	13	39	40
Hispanic	0	*	*	*	*	287	21	24	32	23
Black or African American	1	*	*	*	*	125	21	19	33	27
Native Hawaiian or Other Pacific Islander	1	*	*	*	*	31	13	16	19	52
White	28	14	18	46	21	8459	11	16	38	35
Special Education	7	*	*	*	*	1228	41	24	24	11
Students with a 504 Plan	1	*	*	*	*	55	25	11	45	18
Title I (optional)	11	9	18	64	9	4095	21	21	34	23
Tested with Standard Accommodation	5	*	*	*	*	1875	36	26	27	11
Tested with Non-Standard Accommodation	1	*	*	*	*	14	50	21	29	0
Alternate Assessment	**If a student in your system or school took the CRT-Alternate, please refer to Table III on the CRT-Alternate System or School Summary Report**									
Migrant	1	*	*	*	*	21	24	19	43	14
Gifted/Talented	3	*	*	*	*	640	0	1	16	82
LEP/ELL	3	*	*	*	*	347	54	23	19	4
Former LEP Student	0	*	*	*	*	209	24	28	38	10
LEP Student Enrolled for First Time in a U.S. School	1	**Performance levels are not reported for 1st year LEP students**								
Free/Reduced Lunch	11	18	18	45	18	4690	21	21	35	22

*Less than ten (10) students were assessed

MontCAS CRT

Grade: 05
Spring 2010

Reading

State Summary Report

I. Distribution of Scores

Perf. Level	Scores	State		
		Number	% of Students	% of Students in Cat.
Advanced	298-300	3529	34	55
	295-297	928	9	
	293-294	0	0	
	290-292	881	8	
	287-289	448	4	
Proficient	280-286	730	7	31
	272-279	906	9	
	265-271	765	7	
	257-264	587	6	
	250-256	298	3	
Nearing Proficiency	245-249	252	2	10
	240-244	224	2	
	235-239	200	2	
	230-234	179	2	
	225-229	182	2	
Novice	220-224	70	1	4
	215-219	102	1	
	210-214	98	1	
	205-209	59	1	
	200-204	53	1	

II. Subtest Results

Reading		Possible Points	Average Points Earned
			State
Total Points		60	42
Standards	1. Students construct meaning as they comprehend, interpret, and respond to what they read	21	15
	2. Students apply a range of skills and strategies to read	18	13
	3. Students set goals, monitor, and evaluate their reading progress	This standard is not measurable in a statewide assessment.	
	4. Students select, read, and respond to print and nonprint material for a variety of purposes	9	6
	5. Students gather, analyze, synthesize, and evaluate information from a variety of sources, and communicate their findings in ways appropriate for their purposes and audiences	12	8

CRT Performance Level Descriptors

Advanced (287-300)
This level denotes superior performance.

Proficient (250-286)
This level denotes solid academic performance for each benchmark. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Nearing Proficiency (225-249)
This level denotes that the student has partial mastery or prerequisite knowledge and skills fundamental for proficient work at each benchmark.

Novice (200-224)
This level denotes that the student is beginning to attain the prerequisite knowledge and skills that are fundamental for work at each benchmark.

MontCAS CRT

Confidential

Reading

State Summary Report

Grade: 05
Spring 2010

III. Results for Subgroups of Students

Reporting Category	State				
	Number	% in N	% in NP	% in P	% in A
All Students	10491	4	10	31	55
Gender					
Male	5454	4	11	33	51
Female	5037	3	9	29	60
Ethnicity					
American Indian or Alaska Native	1292	11	25	39	25
Asian	91	3	5	32	59
Hispanic	301	5	13	38	44
Black or African American	140	7	9	35	49
Native Hawaiian or Other Pacific Islander	26	0	4	31	65
White	8641	2	8	30	60
Special Education	1130	18	28	34	19
Students with a 504 Plan	45	7	16	40	38
Title I (optional)	4333	6	16	37	42
Tested with Standard Accommodation	1619	13	28	39	20
Tested with Non-Standard Accommodation	47	28	26	40	6
Alternate Assessment		*	*	*	*
Migrant	22	5	9	55	32
Gifted/Talented	711	0	1	4	95
LEP/ELL	315	28	38	27	7
Former LEP Student	259	5	23	46	26
LEP Student Enrolled for First Time in a U.S. School	1	**Performance levels are not reported for 1st year LEP students**			
Free/Reduced Lunch	4632	6	15	37	41

*Less than ten (10) students were assessed

Appendix P—SAMPLE INTERACTIVE REPORTS

Confidential

Roster and Item-Level Report

Mathematics

System:	**Demonstration District A**
School:	**Demonstration School 1**
Grade:	**04**
Date:	**9/15/2010**

		Released Items																													Total Test Results								
	Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Points Earned by Standard on CRT						Total Points Earned on CRT	Scaled Score	Performance Level
	Content Standard	2	5	2	5	2	6	6	3	7	2	2	7	2	4	3	4	4	7	2	4	7	5	6	6	5	3	2	3	5	4	Standard 2	Standard 3	Standard 4	Standard 5	Standard 6	Standard 7			
	Depth of Knowledge Code	2	2	2	2	2	1	2	2	2	1	3	2	1	2	1	2	1	2	1	1	2	1	2	2	2	2	2	1	2	1									
	Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	SA	CR	MC	MC	MC	MC	MC	MC	MC	SA	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC									
	Correct MC Response	B	A	C	B	D	C	C	A	C			D	C	C	A	B	B	C		B	D	C	D	B	B	B	C	D	C	B									
Name/Student ID	Total Possible Points	1	1	1	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**22**	**8**	**10**	**10**	**8**	**8**	**66**		
AKIN, TEEAZHA	D04100041	D	+	+	+	+	A	+	C	B	1	0	+	+	+	D	+	C	+	1	+	B		+	+	+	+	A	+	D	A	12	4	6	7	7	2	38	255	P
BREWSTER, PARKER	D04100046	+	+	+	A	C	A	+	+	B	1	2	+	D	D	B	+	D		1	A	+	+	C	+	+	+	A	+	A	+		6		8	5	6	40	264	P
BUZBY, ELIJAH	D04100040	C	+	A	+	C	+	+	+	D	0	1	+	+	+	+	+	+	D	0	+	+	B	+	A	D	+	+	+	D	+	9	5	8	6	7	2	37	252	P
CALHOUN, JACOB	D04100042										B	0								0												0	0	0	0	0	0	0	200	N
CAMACHO, ARACELI	D04100025	+	+	+	C	+	+	+	+	D	1	3	+	+	+	+	+	D	+	1	+	+	+	+	+	A	+	+	+	+	+	19	7	7	8	8	7	56	300	A
CLARK, JORDAN	D04100004	+	+	B	+	+	B	+	+	D	1	3	B	B	+	+	+	D	B	0	D	C	A	A	C	+	C	A	C	D	D	12	2	4	7	4	3	32	237	NP
CLINCH, ARIEL	D04100031	+	+	+	C	C	+	+	D	B	0	0	B	B	+	D	D	+	A	0	+	C	B	C	+	D	+	A	B	D	C	4	1	6	4	5	0	20	201	N
COLE, PEYTON		C	+	B	D	B	+	+	+	A	0	1	+	+	+	D	+	D	D	0	D	B	A	+	+	C	+	B	+	+	+	6	7	7	5	8	1	34	243	NP
COLE, TAYLOR	D04100015	+	+	+	+	C	+	+	C	D	0	2	+	+	+	D	+	D	A	0	D	C	B	+	C	+	+	D	+		C	10	5	5	8	4	1	33	240	NP
FATA, WAYNE	D04100007	+	+	A	+	B	D	+	+	B	B	2	C	+	D	C	C	C	D	1	+	B	+	C	+	+	+	D	C		D	11	2	3	6	4	0	26	219	N
HALE, ALEXANDE	D04100038	+	+	+	C	+	+	+	C	+	0	1	+	D	+	+	+	+	+	0	A	B	A	+	+	+	A	D	+	D	+	9	4	8	6	8	3	38	255	P
LEDWELL, SPENCER	D04100024	+	+	+	+	+	A	+	+	D	1	4	+	+	+	+	+	+	+	1	+	+	+	+	+	+	+	D	+	D	+	21	8	10	9	7	7	62	300	A
LOVELESS, LEAH	D04100026	+	+	+	+	C	+	+	+	+	1	1	+	+	+	+	+	+	D	1	C	+	+	+	D	+	+	D	A	A	+	13	5	6	8	6	4	42	266	P
MACAPAGAL, MARK	D04100030	C	D	A	+	+	+	A	+	+	0	1	+	+	+	D	+	D	+	1	+	+	+	+	+	D	+	D	C	D		9	5	5	6	5	7	37	252	P
RIVINIUS, SHAYLAH	D04100035	+	+	+	+	+	A	+	+	D	0	3	+	+	+	+	+	+	D	0	+	+	D	+	+	+	+	A	C	D	+	12	3	7	7	6	5	40	260	P
ROBINSON, AUSTIN	D04100039	A	+	B	+	+	A	+	+	D	0	1	B	+	+	B	+	C	B	1	D	B	+	+	+	C	+	D	+	+	+	9	5	7	8	7	0	36	249	NP
RYEA, AUTUMN	D04100003																																					0		ALT
SAMPO, NATHAN	D04100047	D	B	D	D	+	A	B	+	B	0	1	B	A	+	B	+	D	B	0	A	B	+	A	+	D	+	B	A	D	A	4	2	6	4	3	0	19	200	N
SCHREFFLER, JACOB	D04100027	+	+	+	+	+	+	+	C	+	0	2	+	+	+	+	C	+	+	1	C	+	+	+	+	+	+	D	+	+	+	13	6	6	10	7	8	50	288	P
SILVEIRA, ISAIAH	D04100028	D	+	D	A	+	A	+	B	+	0	3	+	+	+	D	+	D	D	1	+	C	A	C	D	C	+	B	+	D	+	14	2	9	5	3	2	35	246	NP
ZARAGOZA, MELISSA	D04100044	+	B	D	+	+	+	+	+	D	0	1	+	+	+	+	+	C	+	0	D	C	D	+	C	A	A	+	+	D	+	9	5	5	4	5	2	30	231	NP
	Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30									
	Percent Correct/Avg. Score: Group	60	87	53	67	67	60	93	67	33	0.3	1.6	80	87	93	53	80	47	47	0.6	60	47	47	80	73	47	87	20	67	27	73	**11.3**	**4.6**	**6.7**	**6.6**	**6.2**	**3.3**			
	Percent Correct/Avg. Score: School																																							
	Percent Correct/Avg. Score: System																																							
	Percent Correct/Avg. Score: State	69	88	62	71	74	70	78	69	49	0.6	2.1	84	66	80	64	94	53	54	0.5	77	65	54	67	71	59	79	28	81	42	82	**12.7**	**5.3**	**7.1**	**7.4**	**6.0**	**4.6**			

Performance Level Summary

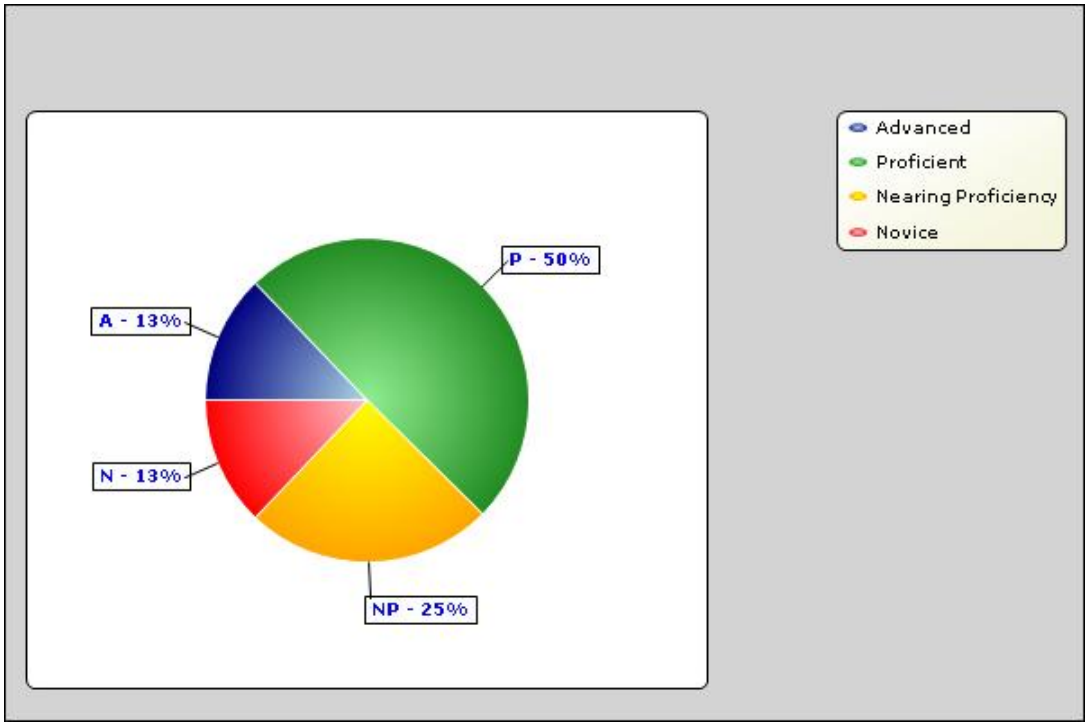
System: Demonstration District A

School: Demonstration School 1

Grade: 04

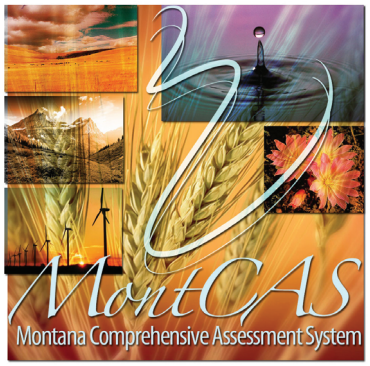
Date: 9/15/2010

Mathematics



Performance Level	Count	Percentage %*
Advanced	2	13
Proficient	8	50
Nearing Proficiency	4	25
Novice	2	13

*Percentages may not total exactly 100% due to applied rounding.



CONFIDENTIAL

Longitudinal Data Report

Student Name
TEEAZHA AKIN

Year	Enrolled Grade	School Name	Administration	Test Name	Content Area	Score	Achievement Level
0910	04	Demonstration School 1	MontCAS CRT	Grade 04 Mathematics	mat	255	Proficient
0910	04	Demonstration School 1	MontCAS CRT	Grade 04 Reading	rea	248	Nearing Proficiency
0910	04	Demonstration School 1	MontCAS CRT	Grade 04 Science	sci	232	Nearing Proficiency

Note: This report returns as many years of data as are available for this student beginning with 06-07.